48

Support Document #23

EPA 540/09-88-109
September 1988

GUIDANCE DOCUMENT FOR CONDUCTING

TERRESTRIAL FIELD STUDIES

Prepared by Edward C. Fite, Larry W. Turner, Norman J. Cook and Clayton Stunkard
Ecological Effects Branch, Hazard Evaluation Division, Office of Pesticide Programs

United States Environmental Protection Agency
Washington, D.C.

ACKNOWLEDGEMENTS

The authors wish to thank those individuals or groups who reviewed and commented on the various drafts of this document. The comments and recommendations received proved helpful and were pertinent to the techniques and methodologies presented. We are especially grateful to the members of the Subpanel of the Science Advisory Panel, Harold Bergman (Chairman), John M. Thomas, Robert K. Ringer, Douglas Robson, Barbara Walton, Stanley Temple and Jerry Schnoor, for providing invaluable comments on the issues presented by Agency staff. With their recommendation we were able to expand and clarify many of the document's sections. We owe a special thanks to Douglas Robson, Professor of Biological Sciences, Biometrics Unit, Cornell University and member of the Subpanel, for providing extensive input and guidance on the statistical portions of the document. Also, we are grateful to Al Vaughn, Richard Lee, Pierre Mineau, Brian Collins, Bill Williams, John M. Emlen, Rick Bennett, Anne Fairbrother, Linda Lyon, Bill Jacobs, Robert Whitmore, Louis Best, John McCarthy, James Gilford, Richard Tucker, Robert McLaughlin, William Gross, Phil Ross, Les Touart, Dan Rieder, James Goodyear, Doug Urban, Henry Craven, Ann Stavola and Margaret Rostker for their helpful comments. We also would like to credit the ASTM Subcommittee of Upland Field Studies which undoubtedly influenced us through discussions at their meetings we attended and Mike Slimak, Chief, Ecological Effects Branch, during the initial development of this document.

We extend a special thanks to Ms Anne Barton, Acting Director of the Hazard Evaluation Division, for her continuing support and direction on this project.

PREFACE

This document is a technical paper intended to provide guidance on how to perform terrestrial field studies, those studies designed to address the potential adverse effects of proposed pesticide use(s) to nontarget wildlife. These studies are presented as outlined in § 71-5 of the Pesticide Assessment Guidelines, Subdivision E - Hazard Evaluation: Wildlife and Aquatic Organisms, EPA-540/9-82-024, October 1982. Such studies represent Tier IV, the most complex of the terrestrial tests presented in Subdivision E. They are required to support those pesticide uses the Agency determines are likely to result in adverse effects to nontarget terrestrial wildlife. Such studies consist of testing performed in the field under actual pesticide use conditions and, generally, they address the potential acute, subacute and/or chronic adverse effects of pesticide residues to nontarget mammals and birds. The effects to birds and mammals are emphasized because the lower-tier Subdivision E tests usually employ these organisms, but effects to other terrestrial organisms, such as amphibians and reptiles are also examined and considered. Terrestrial field studies, as discussed in this paper, are typically multiyear/multisite studies and consist of two levels of tests: a level 1 or screening study, which essentially determines *if* adverse impacts occur to nontarget wildlife under actual pesticide use conditions and a level 2 or definitive study, that *quantifies* those adverse effects identified in the screening study or from other information. Also, the Agency requires that these tests be performed *only* with nonendangered organisms and *only* in areas where impacts to endangered or threatened species will not occur.

As an amplification of § 71-5 Subdivision E, this paper discusses a variety of basic biological research techniques and wildlife investigative methods for use in assessing the effects of pesticides in the field. These methods and techniques are not new, for the majority of them have been used by wildlife biologists, fisheries biologists and game managers for decades. They are presented here, along with adequate references, in order to assist scientists planning to undertake terrestrial field studies. This document is intended to provide guidance (it is not a cookbook or checklist) and will be updated by the Agency as the state of the art for performing these studies advances.

TABLE OF CONTENTS

INTRODUCTION

Data from full scale terrestrial field studies are required by 40 CFR 158.145 on a case-by-case basis to support the registration of an end-use product intended for outdoor application. Because these studies are complex and costly, the Agency requires these tests to evaluate only those products that appear to pose significant risks to nontarget wildlife.

Laboratory tests generally are amenable to a high degree of standardization. In contrast, field study protocols must retain a high degree of flexibility. Variables such as chemical mode of action, use pattern, crop type, method of application and species density and diversity make standardization difficult in field studies. Therefore, Subdivision E, Hazard Evaluation: Wildlife and Aquatic Organisms, of the Pesticide Assessment Guidelines (EPA, 1982) provides only a general outline for field studies. Specific protocols must be developed as needed and submitted to the Agency for review. Regardless of the variability among field studies, several key elements common to most field studies can be identified. This guidance document was prepared to identify and discuss these elements as they pertain to terrestrial vertebrates, and to provide a better understanding of the purpose of field studies.

WHEN REQUIRED

The Federal Insecticide, Fungicide, and Rodenticide Act, as amended (FIFRA, P.L. 92-516), specifies that for a product to be registered or for continued registration, EPA must determine that it will not cause unreasonable adverse effects on the environment. The law further states that the Agency must specify what data are necessary to make this determination, but acquiring that data is the responsibility of those requesting registration or continued registration.

For nontarget wildlife species (i.e., terrestrial vertebrates with emphasis on birds and mammals), the Agency requires a series of tests that are arranged in a hierarchical or tier system, progressing from basic laboratory tests to applied field studies. This tier system, detailed in Subdivision E of the Pesticide Assessment Guidelines (EPA, 1982), provides a means to identify materials that may pose an inordinate risk, and at the same time ensures that the process does not irresponsibly limit use of safe materials. Typically, the initial screen consists of a comparison of results from three avian laboratory toxicity tests (an acute oral LD_{50} and two dietary LC_{50} studies) and one mammalian toxicity test (an acute rat oral LD_{50}) with estimated environmental concentrations (EECs). In addition, when labeling contains directions for using the material under conditions where wild vertebrates may be subject to repeated or continuous exposure to the pesticide, when the material is stable in the environment, or when the material is stored or accumulated in plant or animal tissues, data on avian reproductive effects are required and mammalian reproduction data (from rodent or other mammalian test species) are examined. If environmental concentrations on wildlife food items *are at or below the threshold levels* eliciting a biological response in the avian or mammalian laboratory studies, usually it is assumed that the probability of seriously impacting nontarget species is low. However, for those materials where *environmental concentrations exceed the lowest-effect level eliciting a biological response*, field studies usually are considered.

In assessing the need for field studies for those chemicals whose EECs exceed the lowest-effect level (LEL), a great deal of judgment is required. Several factors, or

appropriate combinations of these factors, need to be considered in addition to the basic laboratory data and EECs. These include:

- The chemical properties of the pesticide (e.g., persistence, conversion to toxic metabolites, retention on food, repellency);

- Intended use pattern (e.g., treated habitats, expected presence of species, including endangered species, extent of use areas, number of applications and treatment intervals);

- Margin between EEC and the LEL;

- Dose/response relationships noted in laboratory tests, including slope of dose-response line, time of mortality or reproductive effects, signs of intoxication and abnormal behavior, and gross pathological changes as noted in gross necropsies.

When the margin between the exposure level and the lowest-effect level is small the likelihood of a decision by the Agency to require an actual field study is small. The other factors mentioned above are seriously considered in this situation where more judgment is needed. Conversely, when the margin between the exposure level and the toxicological effect level is great and begins to approach, for example, the LC_{50}, then the likelihood of a decision by the Agency to require an actual field study is great. However, the final determination of whether a field study will be required is based on the weight of evidence, factoring in all pertinent information. An in-depth discussion of how the Agency determines when terrestrial field studies are required is in Appendix E.

OBJECTIVE OF FIELD STUDIES

The purpose of the field study is to either refute the assumption that risks to wildlife will occur under conditions of actual use of the pesticide and/or to provide some quantification of the risk that may occur. The purpose is twofold because the FIFRA requirement to determine unreasonable adverse effects implies the need for a risk-benefit analysis. Thus if the assumption of risk cannot be refuted, and in order to refine the risk-benefit analysis, field studies should quantify the adverse effects that would occur from actual use of the pesticide.

A study designed to refute hazard is unusual in biological research. Typically, an investigator is more concerned about concluding with a high degree of confidence that an effect occurred, not that it failed to occur. FIFRA specifies that a pesticide is to be registered only if EPA determines it *will not* cause unreasonable adverse effects. While the difference between an objective of "will cause" and "will not cause" may seem trivial, it substantially influences study design and the evaluation of data.

The adverse effects to wildlife that can result from the use of pesticides can be classified as those that affect populations of wildlife and those that affect individuals but not the entire population. Either of these effects may warrant regulatory action, including cancellation or suspension. An adverse effect that results in a reduction in local, regional, or national populations of wildlife species is clearly of great concern. A pesticide that can repeatedly or frequently kill wildlife is also of concern even if these repeated kills may or may not affect long-term populations. The terrestrial field study, accordingly, must be designed to adequately assess both of these types of effects. In most cases, however, the assessment of population effects presents the greatest difficulties, and a study adequate to assess this effect will also assess the degree of risk to individual wildlife. Consequently, throughout this Guidance Document the primary emphasis is on designing a study to assess the risk of a population effect; the study, however, must also be adequate to address the risks to individual organisms.

The field study must be designed to provide data that show whether wildlife species *will not* be affected significantly by a pesticide under normal pesticide use practices. To fully achieve this objective at the population level, one must have detailed knowledge of the population dynamics and varying environmental conditions for each species potentially at risk. The theoretical aspects of population dynamics are well documented in the literature. However, empirical data are available for only a few species (Eberhardt, 1985). A study designed to provide the needed data would include information on age structure, age-specific survival and reproductive rates, and the nature and form of intrinsic and extrinsic regulatory mechanisms. Such a study, when coupled with the influence of pesticide application on these parameters, would require several, if not many years in order to begin to give meaningful results. The cost of obtaining such data could make these studies impractical, if not impossible.

The essential question then is: How can these studies be performed in a practical, economical manner and still provide data that can show that the pesticide under study will not reduce or limit wildlife populations or repeatedly kill wildlife?

One can begin to answer the question by examining the potential influence pesticides can have on wildlife. These effects include:

- Direct poisoning and death by ingestion, dermal exposure, and/or inhalation;
- Sublethal toxic effects indirectly causing death by reducing resistance to other environmental stresses such as diseases, weather, or predators;
- Altered behavior such as abandonment of nests or young, change in parental care, or reduction in food consumption;
- Reduced food resources or alteration of habitat; or
- Lowered productivity through fewer eggs laid, reduced litter size, or reduced fertility.

These effects can manifest themselves in a population through reduced survival and/or lower reproductive success. However, if a field study shows that actual use of a pesticide does not affect survival and/or reproductive success or that only minor changes occur, it would seem reasonable to conclude that the use of the chemical will not significantly impact wildlife. Further, if a field study provides estimates on the magnitude of survival and reproductive effects, one can make reasonable projections on the meaning of the effects to nontarget populations by using available information on the species of concern and basic theories of population dynamics. While less than ideal, field studies that collect information on survival and reproductive effects and use these data to address population parameters should provide a reasonable basis for evaluating potential impacts. This is not to imply that effects on populations are the only concern, however, as indicated previously, a study adequate to assess these effect will also assess the degree of risk to individual wildlife.

This document emphasizes avian and mammalian wildlife. The Agency is also concerned about other terrestrial organisms such as nontarget plants, invertebrates, amphibians, and reptiles. Plants and invertebrates are excluded here from direct study, except as sources of food or pesticides to wildlife. Testing guidelines for nontarget plants and insects are in Subdivisions J and L, respectively. Established protocols, especially for acute and chronic toxicity testing, are available for birds and mammals, but not for reptiles and amphibians. Further, the Agency assumes that "protection" for reptiles and amphibians is provided through the risk assessment process for birds and mammals. Occasionally, however, it may be necessary to adapt these field techniques to apply specifically to reptiles and/or amphibians.

GENERAL APPROACH

Field studies required to support registration have evolved into two types, screening and definitive. The type(s) of study(ies) required depends on the available data on the chemical in question. If the available information is limited to laboratory toxicity data on a limited number of species, coupled with EECs, a screening field study may be appropriate, with the objectives of determining if impacts are occurring and, if so, to what species. If a screening study indicates impacts are occurring, or if other available data suggest that deleterious effects have occurred or are extremely likely, the study design should be quantitative, evaluating the magnitude of the impacts in a definitive study. For some chemicals it may be appropriate to proceed directly to a definitive study without the screening phase. Careful consideration needs to be given to the likelihood of impacts occurring in order to determine which approach to use. In some instances, where there is insufficient information to indicate which species are at risk in the field but available data strongly suggest adverse effects will occur, it may be appropriate for a field study to begin with the general approach of a screening study, followed by a quantitative phase that focuses on the species affected in the screening phase. In certain instances there may be sufficient data and information for the Agency to decide additional testing including field testing is not necessary prior to a regulatory action.

The design of field studies differs between the screening study and the definitive study. If the objective of the study is to determine if impacts are occurring, "pass-fail" methods can evaluate whether or not animals are being stressed by the application. These methods may include carcass searching, residue analysis of species collected on study plots, residue analyses of wildlife food sources found in and adjacent to the area of application, behavioral observations, and enzyme analysis. At the quantitative level (definitive study), the objectives should include estimating the magnitude of acute or secondary mortality caused by the application, the existence and extent of reproductive effects, and the influence of pesticide use on the survival of species of concern. Methods that can be used to address these objectives include mark-recapture, radio telemetry, line transect sampling, nest monitoring, territory mapping, and measuring young to adult ratios.

SAMPLING AND EXPERIMENTAL DESIGN

While examples of acceptable experimental designs are given, it is beyond the purpose of this paper to cover the fundamentals of this topic. Appendix A lists several general and specific references that can provide an in-depth review of this subject. Appendix B provides a general outline for a field study protocol to be submitted to the Agency for review. The following sections generally outline points to be considered in designing screening and definitive field studies. As stated in the introduction, specific protocols must be developed on a case-by-case basis and submitted to the Agency for review.

SCREENING STUDY

OBJECTIVE AND SCOPE

The screening study is designed primarily to demonstrate that the hazard suggested by lower tier laboratory or pen studies does not exist under actual use conditions. The interpretations of screening study results, in most cases, are limited to "effect" versus "no effect" determinations. If the study indicates that the pesticide has caused little or no detectable adverse effect, it may be reasonable to conclude that potential adverse effects are minor. However, when effects are demonstrated, it may be necessary to determine the magnitude of the effects, thus requiring additional testing if pesticide registration or continued registration is still desired. Therefore, when information already available shows that a product has caused adverse effects under normal use conditions, the screening study may be of limited value. In addition, where analysis of laboratory or other data strongly suggest that adverse effects are likely to occur, and are unlikely to be attenuated by field use conditions, it may be appropriate to proceed directly to a definitive field study.

In general, the screening study is limited to addressing the potential for acute toxic effects, such as direct poisoning and death, and sublethal toxic effects potentially affecting behavior and/or survival. In most instances, the screening study would not address chronic effects, such as reduced reproduction, or effects such as changes in density or diversity of populations.

In addition, further laboratory and/or pen studies may be useful prior to proceeding to the field, or may be necessary to interpret results of the field study. For example, additional toxicity data on species that are expected to be exposed from the proposed use pattern may indicate which species are more susceptible to the pesticide, allowing the study to be designed to monitor those species in greater depth as well as to provide insight into field results that show some species were affected more than others. Under such circumstances, additional laboratory studies may be unavoidable. If residue concentrations in resident species are being used to indicate potential problems, the relationship between tissue levels and the dose(s) that cause(s) adverse effects must be estimated. If secondary poisoning is of concern, feeding secondary consumers (held in captivity) prey items collected in the field following the application can be useful to evaluate this potential exposure route. Also, laboratory toxicity tests for secondary consumers coupled with residue analysis of prey items can indicate the potential for secondary poisoning of nontarget species. In designing field studies, the utility of laboratory and/or pen tests should not be neglected, and where appropriate their use is encouraged.

GEOGRAPHIC AREA SELECTION

The selection of geographical areas for evaluating pesticide impacts on wildlife can be difficult particularly for pesticides to be used on crops grown over large and diverse areas. Ideally, studies should be performed in each biogeographic area where the pesticide could be used. While this approach may be practical for uses restricted to localized areas and conditions, many uses (e. g., corn, soybeans, alfalfa) would require an inordinate number of studies in different geographic areas, due to the diversity and variability in wildlife species and habitats involved. To keep the number of geographic areas at a manageable level while still accomplishing the purpose of the field study, geographical area selection should be biased toward situations likely to present the greatest risk. If hazards appear to be low under these conditions, it can be reasonably concluded that impacts under less severe conditions would be minor.

A careful review of the species and habitats in the various geographical areas where the pesticide could be used is necessary to identify the areas of highest concern. A sound

understanding of the biology of the species that are found in association with the potential use sites is essential. Identifying these areas may require an extensive literature review and consultation with experts familiar with the areas and species of concern. The study area selected should be frequented by those species that would have high exposure, based on their feeding or other behavioral aspects. If exposure and fate (e.g., degradation) parameters vary geographically, study area selection also should be biased towards maximizing residues available to wildlife. In some circumstances preliminary monitoring of candidate areas may be necessary to determine which should be selected for detailed study.

STUDY SITE SELECTION

Selection of study sites within each geographic area also is extremely important in designing field studies. Ideally, study sites should be randomly selected throughout the study area. This approach may be practical for some areas such as rangeland or large contiguous crops. However, due to the diversity and variability in wildlife species and habitats in most areas, random selection would require a large number of sites to provide a representative sample. The cost and time requirements of such studies would be unreasonable. To maximize the hazard, the sites selected should have associated species that would be at highest risk from the application, as well as a good diversity of species to serve as indicators for other species not present at that specific location. In addition, the choice of study sites that are as similar as possible in terms of abundance, diversity, and associated habitat will facilitate an analysis of the results.

Under some circumstances, it may be difficult to decide beforehand which species are likely to be at highest risk. In most cases, field surveys of a number of sites may be needed to identify which sites should be selected for detailed study. Even when high risk species can be identified, preliminary surveys may be needed to determine which sites have adequate numbers of the high risk species as well as a good diversity of other species.

In general, study sites should be selected from what is considered to be a "typical" application area, but at the same time, study sites should contain the widest possible diversity and density of wildlife species. Identifying potential study sites may require consultation with experts familiar with the areas where studies are proposed and, as indicated above, preliminary sampling.

In the initial evaluation of potential study sites, "edge effect" may indicate which sites support the larger and more varied wildlife populations. As stated by Aldo Leopold (1933), "The potential density of game of low radius requiring two or more types[1] is, within ordinary limits, proportional to the sum of the type peripheries." If study sites are selected to maximize "edge effect" the potential for high density and diversity should be increased. One quantitative measure of edge and "edge effect" (Giles, 1978) is the distances around individual plant communities in relation to the unit area of the community. Population densities, in general, are positively related to the number of feet of edge per unit area of community. Study sites chosen to maximize the ratio of edge to core may serve to indicate sites with higher densities and diversities of wildlife species.

While this ratio can be helpful in selecting study sites, the other characteristics of edge should not be neglected in screening potential study sites. Density and diversity of wildlife species are also influenced by the variety in the composition and arrangement of the edge component cover types and its width. Also, the interspersion, the plant types and their association with one another, influences densities of wildlife species. The "edge effect" is the sum of all the characteristics of edge and hence each component needs to be

[1] Type - The various segments of an animal's environment used for food, cover, or other requirements.

considered. An agricultural field with a relative high edge to core ratio may not have as high a density and diversity as one with a lower ratio but greater variety, width and interspersion. In general, edge characteristics can be used to screen potential study sites; however, preliminary sampling of prospective study sites will be needed to identify study sites with adequate density and diversity of wildlife species.

NUMBER OF SITES

The number of sites needed can be estimated using the binomial theorem. Briefly, the rationale is that for each study site there are two possible outcomes, either "effect" or "no effect." Trials of this type are known as binomial trials and when repeated the results will approximate a binomial distribution. In this case, to use the binomial theorem, one must first define the expected probabilities that birds or mammals on a site are affected or not affected. Then the probability of the discrete binomial random variable x for n replications can be used to determine the minimum number of sites at a certain level of significance.

As an example for discussion purposes, we will define that a problem exists if some specific mortality rate or level of some other variable occurs on more than 20 percent of the potential application sites. Translated into binomial probabilities, there is a 0.2 probability of a site showing an effect and a 0.8 probability of a site not showing an effect. Therefore, if the results from the field trial show that the number of sites affected is significantly lower than .2n, it can be concluded that potential impacts will be below the stated level of concern.

To calculate the minimum number of sites necessary to show a significant difference between the observed and expected, the following formula for the probability of the binomial random variable x can be used (Walpole and Myers, 1972):

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

Where,

x = number of sites showing effects

n = number of sites

p = probability of a site showing an effect

q = probability of a site not showing an effect

Then, solving for n, when x = 0, i.e.,

$$P(x=0) = \binom{n}{0} q^n$$

Let $P(x=0) = \alpha$, then

$$\alpha = q^n$$

$$\log \alpha = n \log q$$

$$n = \log \alpha + \log q$$

Using this formula, the minimum number of sites can be determined. Continuing with the discussion example of 20 percent occurrence of an effect as a level of concern (i.e., a

0.2 probability of an affected site, a 0.8 probability of a noneffected site, and a 0.05 level of significance), n would be:

$$n = \log 0.05 \div \log 0.8$$

$$n = 13.43$$

Therefore, 14 is the minimum number of sites needed such that the probability is not greater than .05 that all sites surveyed would be unaffected. Or, in other words, if 20 percent of the application sites are actually affected, there is only a 5 percent chance of finding all 14 sites unaffected when n = 14. Moreover, if 20 percent of the application sites are actually affected, we expect to find 1, 2, 3, and 4 sites affected with probabilities of 0.15, 0.25, 0.25, and 0.17, respectively, when n = 14.

Under many circumstances, conducting this number of replications may not be practical. However, as indicated previously, if site selection is biased toward hazard, the number of sites can be reduced. While arguable, it seems logical that if the "worst" cases are sampled, a less stringent level of significance could be accepted. While this must be determined on a case-by-case basis, the Agency believes a minimum acceptable level of significance under worst case conditions is 0.2 rather than 0.05 under "average" or "normal" use conditions. At this level, eight sites showing no effect would be required to conclude at the 0.2 level of significance that the effect occurred on less than 20 percent of the application sites; or there is less than a 20 percent chance that all eight sites will be judged unaffected when n = 8 sites. Under some circumstances, this may not seem adequately protective. It should be noted, however, that based on this same design, it could be concluded that, at the 0.1 level of significance, effects are occurring on less than 30 percent of the application sites, and at the 0.05 level of significance, effects are occurring on less than 40 percent of the application sites. Hence, with eight sites, it could be concluded with a relatively high degree of confidence that effects would occur on less than 40 percent of the application sites. Also, because worst-case study sites were used the Agency could have additional confidence that adverse effects would occur on less than 20 percent of all normal application sites.

However, under some circumstances, particularly if endangered species could be exposed from the proposed use, additional replication may be desirable. Under these conditions a high degree of confidence that an effect was a rare occurrence would be required.[2]

The above calculations are for when x is equal to zero, no effects are observed on any site. A similar approach can be used to estimate the number of sites necessary to show a significant result for a critical value of x greater than zero. Again the formula for the probability of the binomial random variable can be used summing the probabilities of x and all outcomes less than x. Then by using increasing values of n, the number of replications required to show statistical significance may be determined for a given level of significance for individual x values. That is:

$$P(X \leq r) = \sum_{x=0}^{r} \binom{n}{x} p^x q^{n-x}$$

[2] Under no circumstances should field studies on pesticides be conducted in areas where endangered species could be exposed.

The minimum value of n occurs when

$$P(X \leq r) = \alpha \text{ level}$$

Continuing the previous example, Table 1 gives the results for $x \leq 1$ and $x \leq 2$ for the previously defined acceptable occurrence level of effect (i.e., a 0.2 probability of an affected site, a 0.8 probability of nonaffected site). From the table, the minimum number of sites needed when the critical value for x is set at 1, to conclude that (at a 0.2 level of significance) effects are occurring below levels of concern is 14. If $x \leq 2$, 21 are needed in order to have an equivalent criterion. As can be seen, as x (the number of sites with effects) increases, the number of sites required to show a statistical significance becomes inordinately large.

Table 1.

Probabilities for binomial random variable with p equal to .2 for $x \leq 1$ and $x \leq 2$ as a function of the number of sites (N).

N	P ($x \leq 1$)	P ($x \leq 2$)
8	0.5033	0.7969
9	0.4362	0.7382
10	0.3758	0.6778
11	0.3321	0.6174
12	0.2749	0.5583
13	0.2336	0.5017
14	0.1979	0.4481
15	0.1671	0.3980
16	0.1407	0.3518
17	0.1182	0.3096
18	0.0991	0.2713
19	0.0827	0.2369
20	0.0692	0.2061
21	0.0576	0.1787

When the probability of an affected site is 0.2, application of the rule of "zero observed affected sites" results in a declaration of "no effect" 16.8 and 13.4 percent of the time for samples of size 8 and 9, respectively. It also results in a declaration of "no effect" 43.1 and 38.7 percent of the time for samples of size 8 and 9, respectively, when the probability of an affected site is 0.1, a value less than the criterion probability.

Under any condition, it is extremely important with the binomial approach to define the critical or threshold level for an effect, and to be sure that the methods used are sensitive enough to detect an effect should one occur. These assessments depend upon the species potentially at risk as well as the parameter being sampled. It should be noted that the

measure of effect is not limited to mortality. Other parameters, such as residue or enzyme levels, can be used. Whatever parameter(s) is (are) used, defining the criteria level for an effect is extremely important, and when designing studies this issue should be considered carefully.

Using this approach, control (reference) sites are not an absolute necessity. While the Agency encourages their use, in some cases the additional information gained from the control sites for a screening study may not justify the additional effort required. In most instances, control sites would serve to protect from erroneously attributing effects due to other causes to the pesticide. However, for most chemicals, this can be avoided by employing methods, such as residue analysis and/or cholinesterase inhibition tests, that can be used to indicate if the pesticide contributed to the observed effect. Further, studies have shown that it is a relatively rare event to locate dead or sick animals in the wild except under unusual conditions (Heinz *et al.*, 1979). Thus it is unlikely to find dead animals that were killed by something other than the pesticide being tested.

Nevertheless, in some instances, particularly when reliable methods to confirm the cause of effect are not available, controls may be necessary. In these cases the above binomial design can be modified to a paired plot binomial design, with a treatment plot and a comparable control plot for each study site within an area. Then, as above, when critical levels of effect and occurrence are defined, the binomial theorem can be used for sample size determination, which gives eight site pairs (16 paired plots) showing less than a defined difference between plots to conclude at the 0.2 level of significance that the effect occurred on less than 20 percent of the application sites. Alternatively, a quantitative difference or, preferably, ratio of treated to control responses could be used to test for a treatment effect on each of the measured response variables. (This is discussed further in the section on experimental design for definitive studies, page 19.)

SIZE OF STUDY SITES

Study sites must be large enough to provide adequate samples. The size is dependent on the methods used, the sensitivity required, and the density and diversity of species and their ranges. In some cases, particularly with slow-acting poisons or where species at high risk have relatively large home ranges, areas several times larger than the treatment area may need to be examined. In some circumstances, several fields in an area may be included in a single study site to account for wide-ranging species or lower densities. Except in the unusual circumstance where "fields" are extremely large (e.g., forested and range areas), the study site should never be less than an individual field and the surrounding area. The nature of the surrounding area is discussed further under individual methods. Another consideration is the distance between study sites. In general, sites should be separated adequately to ensure independence, which is dependent mainly on the range of the species that could be exposed.

CHEMICAL APPLICATION

In planning field studies consideration should be given to application rates and methods. In general, the test conditions should resemble the conditions likely to be encountered under actual use of the product. In most instances the pesticide should be applied at maximum use rates and frequencies specified on the label.

If more than one application method is specified on the label, the method that maximizes exposure of nontarget species should be used. This evaluation should relate wildlife utilization of the area to exposure. For example, if the crop is one that is used by avian species as preferred nesting areas, feeding areas, or cover, then ground application may be the method that maximizes exposure. However, if it is a crop with

low utilization by wildlife species, but with high utilization of its edges, aerial application where drift could increase exposure may be more appropriate. In any case, the method of application used must be consistent with the label.

In addition, the equipment used may influence potential exposure of nontarget species. There is a diversity of types of farming equipment that, depending on the particular use pattern involved, could influence exposure. For example, for pesticides applied in-furrow at planting there are several types of covering devices employed on seeders, such as drag chains, drag bars, scraper blades, steel presswheels, etc., in which the efficiency may vary for covering the pesticide. In general, one must evaluate the various equipment normally used for the particular pesticide application to estimate the potential influence of equipment choice on exposure. In some instances, preliminary tests may be required to estimate which method and equipment poses the highest exposure.

METHODS

This section provides a general outline of methods appropriate for use in a screening field study and indicates some of their limitations. The methods described have been found to be most useful. However, we emphasize that a screening study is not limited to these methods. If other methods are more appropriate, their use is encouraged. Because procedures should be adapted to specific situations, the outlines presented should not be interpreted as strict protocols. Normally, different methods will be combined to evaluate potential impacts. Due to the indefinite number of variables and the unpredictability of wild animals, even normally reliable procedures can sometimes prove inadequate.

Essentially, the methods used in a screening study address exposure by monitoring overt signs of toxicity such as mortality or behavioral modifications, or through evaluating parameters that indicate animals are under stress, such as residue concentrations in tissues or degree of enzyme inhibition. Measurements of density and diversity of species are needed to aid in evaluating the results. The following methods can be useful for screening studies.

Carcass Searches

Searching for dead or moribund wildlife has been a basic method used in field studies to evaluate the impact of pesticides on nontarget species. Carcass searches can roughly indicate the magnitude of kills when adequate areas are searched and the reliability of the search is documented. This latter point is extremely important. Rosene and Lay (1963) indicated that finding even a few dead animals suggests that there has been considerable mortality; failure to find carcasses is poor evidence that no mortality has occurred. The reliability of the search is based upon the percentage of carcasses recovered by searchers and the rate of disappearance. By knowing the reliability, the meaning of the failure to find carcasses can be assessed and the extent of the kill estimated.

Finding dead animals is seldom easy, even if every animal on a site is killed. For example, three breeding pairs of small birds per acre is considered a large population (Heinz *et al.*, 1979), and under average cover conditions, a small bird is difficult to detect. Small mammals may be more abundant but, due to their typically secretive habits, they are more likely to die under cover and be even more difficult to find than birds. Carcass searching specifically for mammals should be attempted only when cover conditions permit a reasonable search efficiency. However, any vertebrate carcasses found should be collected, even if the search is oriented primarily to one taxon.

Because the results may be biased by scavenging and failure to find carcasses, the sensitivity of this procedure should be determined. Under conditions of heavy cover and/or high scavenger removal, other methods may be more appropriate.

There are no standard procedures for carcass searches. Appendix C outlines practices that have been used typically and should be considered in designing searches.

Radio Telemetry

Radio telemetry has been found to be extremely useful for monitoring mortality and other impacts caused by pesticide exposure of wildlife. Advances in miniaturizing electronic equipment over the last 15 years have made it feasible to track most vertebrate animals. Transmitters have been developed that weigh a few grams and have been used to track species as small as mice. Cochran's (1980) excellent summary of this technique provides additional details.

Radio telemetry has the advantages of providing information on the fate of individual animals following a pesticide application and of facilitating carcass recovery for determining the cause of death. Although the initial cost of this technique may be more than for other methods, the increase in information obtained under some circumstances can more than justify the cost. The method is particularly useful with less common or wide-ranging species.

In addition to mortality, radio telemetry can be used to monitor behavioral modification as well as physiological changes. Automatic radio-tracking systems permit continual surveillance of the location of animals (Cochran, 1980), which could be used to provide insight into behavioral changes such as nest abandonment, desertion of young, or decreases in activities such as flying or feeding. Radio telemetry equipment is also available for the transmission of physiological data such as heart rates or breathing rates (Moen, 1973).

While this technique can provide very useful information on impacts of pesticides to wildlife, other points need to be considered in addition to cost. Capturing animals alive and unharmed requires more time, skill, and motivation than one might expect. For the method to be consistently successful, the investigator must be thoroughly familiar with the habits of the species under study and with the various capture methods that can be used. Even for the most experienced investigator, adequate sample sizes can be difficult to obtain under some conditions.

Adequate sample size is very important. The binomial theorem can be used to estimate minimum sample size per site, if the question is limited to mortality. Briefly stated, to be sure that nontarget species are not being affected by environmental concentrations greater than, for example, an LC_{20}, the expected binomial probabilities would be 0.2 for mortality and 0.8 for nonmortality. Depending on the level of significance, 8 ($\alpha = 0.2$) to 14 ($\alpha = 0.05$) individuals would need to be monitored per site (see section on "number of sites" for further details on these calculations). However, since the LC_{20} may differ between species, 8 to 14 individuals would be required for each species, unless laboratory tests have documented relative species sensitivity. Further complications can arise if the radio-tagged animals leave the area or if the movements of individuals limit their exposure. If these complications occur at relatively low rates, a few additional radio-tagged animals may be sufficient to overcome these problems.

Tests of Cholinesterase Inhibition

Measuring cholinesterase (ChE) concentrations in animal tissues has been found to be a very useful field technique for evaluating exposure of nontarget animals to ChE inhibiting chemical (Heinz *et al.*, 1979; Hill and Fleming, 1982). These chemicals, including organophosphates and carbamates, affect the synaptic transmission in the cholinergic parts of the nervous system by binding to the active site of acetylcholinesterase (AChE), which normally hydrolyzes the neurotransmitter acetylcholine. Thus, ChE inhibitors permit excessive acetylcholine accumulation at synapses thereby inhibiting the normal cessation of nerve impulses (O'Brien, 1967; Corbett, 1974).

The depression of AChE activity, when measured and compared to controls, can indicate the degree to which an animal is affected. Brain ChE depression of > 50 percent in birds has been found sufficient to assume that death is pesticide related (Ludke *et al.*, 1975); depressions of more than 70 percent are often found in dead birds poisoned by these chemicals (Bunyan *et al.*, 1968a; Bunyan *et al.*, 1968b; Shellenberger *et al.*, 1970), although some individual birds with less than 50 percent inhibition may die (Ludke *et al.*, 1975; Bunyan *et al.*, 1968b). A 20 percent depression of brain ChE has been suggested as an indication of exposure (Ludke *et al.*, 1975). ChE concentrations in blood can also be used to indicate exposure, avoiding the necessity of sacrificing the animal. However, blood ChE concentrations are influenced more by environmental and physiological factors than are brain ChE concentrations. Because ChE activity varies among species, the degree of depression must be based on an estimated normal value for concurrently tested controls of the species potentially at risk. Because of this difference between species, each case must be considered unique (Hill and Fleming, 1982).

Although there are several colorimetric methods for determining ChE activity, the general methods are similar. Brain tissues (or blood samples) are taken and analyzed for ChE concentrations. Comparisons are then made between pre- and posttreatment and between treated and untreated areas. It is important to ensure that "untreated" controls have not been exposed to any ChE inhibitors. It also should be noted that, at the present time, absolute enzyme levels in the literature are derived from various different, although similar, methods and are reported in different ways. For example, Ludke *et al.* (1975) used a modification of the Ellman *et al.* (1961) method and reported results of ChE activity as nanomoles of acetylthiocholine iodide hydrolyzed/ minute/mg of protein, whereas Bunyan *et al.* (1968a) used their own colorimetric method (in addition to a pH change method) and reported the results as micromoles of acetylcholine hydrolyzed/ hour/mg of protein. Therefore, without a tightly standardized method, it is necessary to use concurrent controls of the same species obtained from the general vicinity (but untreated) of the exposed birds, rather than literature values. Because of the greater variation in plasma ChE levels than for brain, more controls are necessary to evaluate blood samples.

Tests for ChE activity can be used to help confirm cause of death and monitor levels of exposure. In the latter case, 5 to 10 individuals of each species are collected before treatment and at periodic intervals following treatment. Mean inhibition of 20 percent or more is considered an indication of exposure to a ChE inhibitor. Confirmation of cause of death may be determined by analyzing brain tissue from wildlife found dead following treatment and comparing the activity with controls. Inhibition of 50 percent or more is considered strongly presumptive evidence that mortality was caused by a ChE-inhibiting compound. The cause-effect relationship can be further supported by chemical analysis of the contents of the digestive tract or other tissues for the chemical in question.

For this technique to provide accurate information, prompt collection and proper preservation of specimens are essential. ChE concentrations in tissues are influenced by time since death, ambient temperatures, and whether or not "reversible" ChE inhibitors are being investigated. Therefore, the response of postmortem brain ChE to ambient conditions can seriously affect diagnosis of antiChE poisoning. Samples must be collected shortly after death and frozen immediately to halt changes in tissue or enzyme-inhibitor complexes (Hill and Fleming, 1982).

Hill and Fleming (1982) have reviewed a technique for field monitoring and diagnosis of acute poisoning of avian species, discussing sample collection, sample numbers, preservation procedures, and sources of error. Their publication is recommended for review for additional details.

Residue Analysis

Residue analyses of wildlife food sources provide information about the level and duration of pesticide exposure. Residue analysis of animal tissues also can indicate actual exposure levels. If the relationship between tissue concentrations and toxic effects is known for the species in question, residue analyses can provide a measure of the degree to which the animals were affected. For this application of residue analyses though, laboratory trials are necessary to establish the relationship between residue levels and toxicity. In addition to death, these laboratory trials should include such signs as anorexia, asthenia, asynergy, or ataxia. For chemicals that are readily metabolized by vertebrates, residue analysis may not be appropriate for diagnostic purposes. With many pesticides, it will be necessary to analyze also for residues of active metabolites.

For determining residues on wildlife food sources, the investigator should collect samples of insects, seeds, leafy parts of plants, etc., immediately after pesticide application and at periods thereafter. Samples should be analyzed for the chemical to determine potential exposure rate and duration. The application method needs to be considered in determining where to take samples. If drift is likely, samples should be taken from habitats surrounding the treatment sites as well as in the treated fields. Because analysis can be costly, the investigator should consider carefully the number of samples necessary to provide adequate data. Where feasible, samples from different locations within a site should not be pooled. Separate analysis of samples can provide data on the range and variability of exposure as well as mean levels.

When residue analysis is used to evaluate exposure in nontarget animals, the tissues selected for analysis differ depending on the purpose. Heinz *et al.* (1979) indicated that residues in brains of birds and mammals can be used to determine if death is pesticide related for many chemicals. Sublethal exposure, they believe, is judged better from residues in other tissues. Therefore, Heinz *et al.* (1979) propose analyses of whole body homogenates to quantify the body burden of a pesticide. If this is not feasible, they suggest analyzing muscle tissue, because muscle residues reflect body burden more nearly than those of any other tissue, and the amount of muscle tissue is not unduly large. For persistent chemicals, Heinz *et al.* (1979) suggest that residues in liver and fat tissues could be misleading for determining acute body burdens. Liver is a processing organ and its residue level largely represents current availability of the chemical. Residues in fat are greatly affected by changes in the amount of body fat, and are undependable indicators of body burden of the chemical. However, for some chemicals, liver, fat or other tissues may be good qualitative indicators that exposure did occur. In general, laboratory trials or data gathered in metabolism or other studies may be necessary to determine which tissues can provide the most useful information. Residue analysis of eggs taken from nests in treatment areas can indicate the degree of contamination that a treatment has caused, as well as possible reproductive effects of the treatment.

Two approaches may be used to determine the number of samples to be collected. Frequently, residue samples will be collected to establish a mean value and confidence limits. To determine the number of samples necessary to collect, it is necessary to estimate the standard deviation and to set arbitrarily a limit from the mean value that is acceptable. Although the mean value does not need to be estimated, it is also necessary to have some idea of the mean so that the standard deviation can be estimated and the limit can be set. The formula for the number of samples, as presented by Snedecor and Cochran (1967), is: $n = 4\sigma^2 \div L^2$ for 95 percent probability, where σ is the standard deviation and L is the allowable limit around the mean. For example, if one wants to know the residue concentrations on vegetation within $\pm$ 10 ppm and estimates a standard deviation of 20 ppm, then $n = 4(20)^2 \div (10)^2$ or 16 samples are required to have a 95 percent probability that the sample mean value will be within $\pm$ 10 ppm of the true mean.

In some situations, there may be little information useful for estimating the standard deviation, or the standard deviation may be rather large, thus requiring a very large

sample size. For some types of samples, such as residues in nontarget wildlife carcasses, the sample size cannot be increased to permit more precision. The mean value of a parameter certainly has utility, but it also is very important to establish confidence limits around the mean. In general, the Agency will use the 95 percent confidence limits (usually the upper boundary, as in the case of residues) in the assessment of the data. This approach will substantially reduce the impact of outliers but will still incorporate the range of reasonable values into the assessment. In addition, the use of confidence limits reduces the necessity for taking a large number of samples. Of course, the width of the confidence intervals decreases with increasing sample sizes; so an investigator should take as large a sample as feasible.

Since the sample size will nearly always be less than 30, the calculation of confidence limits should be based on Student's t-distribution. The t values are derived from tables available in most statistics books, and the 95 percent confidence limits are:

$$\bar{x} \pm (t_{.05}) (s \div \sqrt{n})$$

where s is the standard deviation estimated from the sample of size n.

Alternatively, the binomial approach may be used for determining if residues, typically in collection of live nontarget animals, exceed a particular threshold value that indicates an effect. The required sample size is the same as presented for the binomial approach in determining the number of study sites; specifically, in the preceding example a minimum of 8 samples with none exceeding the threshold value or 14 samples with one or none exceeding the threshold value indicates "no effect" at p = 0.2 in 20 percent of the samples. This approach requires the establishment of threshold values which are determined on a case-by-case basis. In general, residues reflecting an LC_{20} level of exposure would seem to be a maximum acceptable effect concentration for a screening study. Ideally, for each species analyzed for residues, an LC_{50} would be determined in the laboratory. Then a group of animals would be exposed to an LC_{20} concentration to determine the mean threshold concentration of residues. Since this approach is impractical for a screening study, it is suggested that the mean residue concentration in bobwhite and/or mallards exposed to an LC_{20} dietary concentration would provide an indication of threshold levels.

The number and timing of collection periods must be considered and should be based on the persistence of the specific chemical under study. Where persistence in the field has not been adequately determined, it may be necessary to sample at regular intervals (e.g., days 0, 1, 3, 7, 14, 28, 56) to provide data on degradation rates.

Behavioral Observations

Observations of behavior sometimes can be an extremely important indicator of treatment effects. Such observations might include characteristic signs of toxicity or behavioral changes seen in test animals exposed to the pesticide in the laboratory. Other abnormal behavior (e.g., territorial males abruptly ceasing singing, birds not feeding, reduced avoidance of humans) also may be important.

Density and Diversity Estimates

It is necessary to know the number of individuals and variety of species on and around a study site in order to indicate which species could have been exposed and to aid in evaluating the significance of mortalities or other findings. In addition, preliminary information on density and diversity is necessary for site selection and to determine the size of study sites. Under some circumstances, comparisons of density estimates between treatment and control sites, or between before and after treatments, may be used to indicate pesticide impacts. In general, the usefulness of these comparisons is limited in a screening study due to the relatively small acreage involved. If mortality occurs, replacement from outside is likely to be so rapid that losses are replaced before censuses

are completed. Seasonal changes, such as migration, molt, or incubation, that can affect real or apparent densities, also must be considered.

Several techniques may be used to estimate the density and diversity of wildlife species, including counts of animal signs, catch per unit effort, mark-recapture, and line transect sampling. (Appendix A provides references on the various techniques available.) Although the methods selected depend on the species of concern, for the screening field test line transect methods are likely to be the most useful for birds.

The major advantage of line transect sampling is that it is relatively easy to use in the field once a proper sample of lines has been chosen. However, line transect sampling is not applicable to all species, particularly those that are not easily observed. Individuals using line transects must be extremely competent in species identification.

In the line transect method, an observer walks a distance (L) across an area in nonintersecting and nonoverlapping lines, counting the number of animals sighted and/or heard (N), and recording one or more of the following statistics at the time of first observation:

- Radial distance from observer to animal;
- Right-angle distance from the animal sighted or heard to the path of the observer; or,
- Angle of sighting from the observer's path to the point at which the animal was first sighted or heard.

Although the field procedures are simple, they must be understood adequately and implemented well to obtain good estimates of density (Burnham *et al.*, 1980).

Burnham *et al.* (1980) provide a thorough review of the theory and design of line transect sampling. This monograph should be reviewed for details along with other references listed in Appendix A.

For mammals, density and diversity estimates from capture data may be the most practical for a screening study. There are several ways of estimating the populations from capture data, some relatively simple, that may provide adequate information for a screening study. Davis and Winstead (1980), as well as other references listed in Appendix A, review the various methods available, explaining their advantages and disadvantages.

INTERPRETATION OF RESULTS

The numerous variables involved in field studies makes a meaningful discussion of the interpretation of results somewhat tenuous, particularly with the almost inexhaustible array of results that could occur. Each study must be considered unique and therefore will require a case-by-case analysis that incorporates not only the actual study but other relevant information that is available. There are a few points that can be discussed however, that may be helpful when designing studies.

In general, the results of the screening field study should provide information on acute poisoning and potential sublethal effects as suggested by enzyme, residue or other measurements. In addition, information will have been developed on the density and diversity of species on the study sites as well as the sensitivity of the methods used. If no effects are detected, assuming that the methods used were adequate to detect levels of concern and that the species on the study site represent a good cross-section of the nontarget species expected to be at risk, the potential hazard indicated by lower tier tests is refuted. Unless other hazards (e.g., reproduction) are still of concern, additional tests would not normally be necessary. However, if an effect is detected on one or more study

sites at rates equal to or greater than concern levels, the hazard has not been refuted and additional tests may be necessary.

In interpreting if an effect has occurred in the context of the binomial approach, care must be employed not to assume a level of precision in results that does not exist. For some methods used in these studies, due to the inherent variability in the data collected to estimate the level of impact, particularly when minimum sample sizes and areas are used, most detectable effects will exceed the concern level. In some instances in interpreting results it may be appropriate to use confidence limits of data collected (or another measure of dispersion) to evaluate if concern levels are exceeded. For example when density estimates are used to estimate percent mortality using the number of dead animals found during carcass searching, the upper and lower confidence limits of the density estimate may be more appropriate than the average, particularly when variability of the density estimate is high. Whatever method is used, when effects are detected that exceed concern levels they will be put into perspective in the context of the entire study as well as other available information to determine if or what additional data are needed. A "no pass" result does not necessarily mean that definitive field testing is automatically required.

For example, a test may be run in an area where a species is abundant, yet on a specific study site their numbers may be sufficiently small that a single death exceeds the level of concern on that site. Statistically, such a finding would indicate that the study did not "pass" according to the binomial approach, and this would be the preliminary interpretation. However, if the other sites had an adequate number of this species as well as other species expected to be at risk and no other signs of impacts are observed, the implications of the mortality would seem minor. On the other hand, if diversity of species were extremely limited, it would have greater significance. In other situations where the one dead bird is of a species with small numbers on most sites, but density and diversity of other species is representative of nontargets expected to be at risk, another screening study that looks at the species in which the effect was detected may be appropriate. Conversely, a screening study showing that there is appreciable mortality on most study sites may be sufficient for the Agency to consider regulatory action.

In summary, the interpretation of results will go beyond the statistical evaluation since the Agency must consider all the factors and circumstances peculiar to each test and site. The biological interpretation of results is, and probably always will be, a matter of scientific judgment based upon the best available data. In general, the judgmental aspects of biological interpretation are more important for definitive studies than for screening studies. Nevertheless, biological considerations often will be relevant to screening studies. Study conclusions must integrate that which is biologically significant with that which is statistically significant.

Another consideration in the interpretation of results of a field study is the attribution of effects to the pesticide being studied. A well-designed study will include appropriate techniques to determine if an effect is caused by a pesticide. In the absence of such techniques, the Agency has no choice but to consider that any effects were as a result of the pesticide use. As an example, measurement of ChE levels can provide information, since it is generally accepted that inhibition of 20 percent indicates exposure and inhibition of 50 percent or more indicates, in birds, that mortality is due to an inhibitor (Ludke *et al.*, 1975). If the test chemical is the only cholinesterase inhibitor used in the vicinity of the study site, it can be reasonably assumed that a mortality associated with 60 percent ChE inhibition is due to the test chemical. However, if other ChE inhibitors are used near the site, additional information, such as residue measurements, may be necessary to attribute death to the specific ChE inhibitor being tested.

Appendix D provides further discussion and examples of what is involved in planning and conducting a screening study.

DEFINITIVE STUDY

OBJECTIVE AND SCOPE

The definitive study is a relatively detailed investigation designed to quantify the magnitude of impacts identified in a screening study or from other information. In contrast to the screening study, which monitors mainly the proportion of the local population that is expected to be exposed, the definitive field study examines a sample of the entire local population in the treated area. Although a definitive study *may* be done when laboratory studies indicate a high potential for field mortality, it is more likely to be requested when there is evidence that actual field mortality has occurred, as in a screening study, or where reproductive effects are being investigated. The objectives of the definitive study are:

- To quantify the magnitude of acute mortality caused by the application;
- To determine the existence and extent of reproductive impairment in nontarget species from the application; and
- To determine the extent to which survival is influenced.

Due to the intense effort and time required to estimate these parameters, the definitive study should be limited to one or a few species believed to be at the highest risk. If it can be shown that minimal (as defined at the onset of the study) or no changes occur in study parameters to high risk species, there is likely to be minimal potential for adversely affecting other presumably low risk species from use of the pesticide in question.

The definitive study, in addition to estimating the magnitude of effects of acute toxicants, also can be applied to estimating the magnitude of chronic or reproductive effects. Although we have emphasized chemicals that are acutely toxic, with few exceptions the discussion is applicable to chemicals that cause chronic effects.

In general, the definitive study will provide limited insight into whether or not effects are within the limits of compensation for the species of concern. However, using the data collected in these studies, coupled with available information on the species of concern and basic theories of population dynamics, the meaning of the observed effects on the species can be evaluated.

SAMPLING AND EXPERIMENTAL DESIGN

As indicated previously, the principles of statistical design of studies are well documented and it is beyond the scope of this document to cover the fundamentals of this topic. However, there are a few points on this topic that warrant discussion relative to the definitive study.

In the design of field studies, one must carefully consider what constitutes a sampling unit. Eberhardt (1978) points out that special problems are faced in designing experiments on wild animal populations. Study sites must be large in order to limit the influence of boundary effects, such as movements into and out of the area. Large study sites can be very expensive both in terms of actually applying the experimental treatment and in the assessment of results. Eberhardt also states that numerous observations, even a full year of data, on a single study site may result in very sound values for that site, but do not provide a basis for inferences to other sites. Hurlbert (1984) has discussed the problems associated with field studies where there was no replication or replicates were not statistically independent, which he terms pseudoreplication. Of the field studies he

evaluated, 48 percent of those applying inferential statistics had pseudoreplication.

According to Eberhardt (1978), lack of replication seems to be based on the mistaken assumption that variances based on subsampling of sites (intrasite variability) are suitable bases for comparing treatment effects (intersite variability). This, he believes, is not a valid basis for a statistical test, because it is the variance of sites that are treated alike that is relevant to a test of treatment differences. Although subsampling of sites may be necessary to collect the data, it is the difference between sites that is important for analysis.

An important point to consider in designing a definitive study is to be sure that the study will detect a substantial impact when, in fact, it occurs. In statistical terms this concept is referred to as the power of the test. Experience with "classical" experimental designs with random assignment of experimental "treatment" and "controls," has shown that the probability of a Type II error is generally high (unless very large numbers of replicates are available). Eberhardt (1978) indicates that, all too often in field studies on impacts to wildlife, either by default or lack of understanding, there is only a 50 percent chance of detecting an effect, which he likens to settling the issue by flipping a coin and doing no field study whatsoever. Since a definitive study is carried out under the assumption that effects will occur, the Agency believes minimizing Type II errors is extremely important.

As suggested above, the more generally used experimental designs require inordinately large sample sizes to obtain small Type II errors. For example, based on a coefficient of variation of 50 percent (a relatively homogeneous sample for the kinds of data collected in field studies; Eberhardt, 1976), a 20 percent minimum detectable difference between means, a Type II error of 0.2 and a Type I error of 0.05, the number of replications required can be estimated as:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \, (CV)^2 \, [1+(1-\delta)^2]}{\delta^2}$$

where

n = number of replications

$Z_{1-\alpha}$ and $Z_{1-\beta}$ are critical values of the unit normal distribution

CV = coefficient of variation

δ = detectable mean difference expressed as the proportion of the control group mean, i.e.,

$$\delta = (\mu_1 - \mu_2) \div \mu_1$$

For the above example:

$$n = \frac{(1.65 + 0.84)^2 \, (0.5)^2 \, [1 + (1-.2)^2]}{(.2)^2}$$

n = 63.6 replicates

Thus, for the above parameter, 64 replicates for both control and treatment groups, or 128 total study plots, are required to detect a 20 percent difference between treatments

and controls with an 80 percent chance of being sure to detect a real difference at a .05 level of significance.

With more sophisticated designs, the number of replicates can be reduced under some circumstances and still meet the Agency's aspiration to limit the probability of a Type II error to 0.2 with a detectable difference of 20 to 25 percent. For example, a paired plot design can be used, substantially reducing the number of replicates required. Pairing serves to reduce the effective coefficient of variation by reducing the variation attributable to experimental error. The lower coefficient of variation reduces the number of replicates. Then a quantitative difference or, preferably, a ratio of treated to the total of treated and control responses, can be analyzed statistically to test for a treatment effect on the measured response variables (SAP, 1987).

The logic of using paired plots is that, while no two areas are ever exactly alike, two areas that are not widely separated in space are ordinarily subjected to much the same climatic factors, have populations with about the same genetic makeup, and generally the two populations can be expected to follow much the same trend over time, apart from a pesticide effect (Eberhardt, 1976). Then, if all plots are approximately equal in area and habitat and population densities between pairs are similar, we are postulating that when no pesticide impacts occur, the mean ratio of treatment to treatment plus control will equal one-half. Then a t-test or an exact randomization test (Edgington, 1980) may be applied to test whether the average number of survivors on the treated plots is equal to the average number of survivors on controls.

The number of pairs required can be estimated using the following formula[1]:

$$n = (Z_{1-\alpha} + Z_{1-\beta})^2 \quad \frac{4q_1}{p_1^2 (1+q_1) \bar{c}}$$

where,

n = number of paired plots

$Z_{1-\alpha}$ and $Z_{1-\beta}$ are critical Z scores

q_1 = survival ratio

p_1 = mortality ratio

$\bar{c}$ = mean number of survivors on control plots

Therefore, at an 80 percent assurance of detecting a treatment-induced impact of 20 percent or greater at a 0.05 level of significance if $\bar{c} = 28$,

$$n = (1.65 + 0.84)^2 \quad \frac{4(.8)}{(.2)^2 (1 + .8)\ 28}$$

$$n = 9.84$$

Thus, 10 pairs of plots (20 total) with a mean of 28 individuals per plot would be needed. Increasing the mean number of individuals per plot ($\bar{c}$), causes a reduction in n.

In some field situations, pairing may not be feasible. In these situations, other designs would be more appropriate or less rigorous design may have to be used. However, in

[1] See Appendix F for development of this formula.

planning field studies, one must be careful to consider the power of the study design to determine the limitations of the study. Studies with adequate replication are highly preferred to support registration; the use of less replication will not necessarily render the study inadequate. However, what is objectionable is to use a study with low power to imply no biological damage, when the study was not capable of detecting it if it occurred. In cases where large numbers of replicates are impractical, subjective and biological knowledge should be used in a decision process to decide if there was a treatment effect. In most instances, it is highly advisable to involve statisticians or biometricians who are familiar with this kind of field study in the planning and analysis phase of the field work to avoid costly technical errors.

STUDY AREA AND SITE SELECTION

Selection of geographical areas and study sites within the areas for the definitive test generally requires the same considerations as for a screening study. For the definitive study, however, the selected areas and study sites must have adequate populations of the species of concern. Obviously, the crop of concern must be grown on a representative portion of the area. Also, consideration needs to be given to whether the target pest species will be present. If it is not, one must consider what influence its absence may have on potential results. For example, if the pest is a major food source for nontarget species, its absence could significantly influence results. Finally, the potential variation in populations of concern over the geographical area(s) selected should be considered. It may be difficult to find sites that are sufficiently similar to provide paired plots, which limits the coefficient of variation so that the desired sensitivity can be achieved.

NUMBER AND SIZE OF SITES

As suggested in the section on study design for the definitive test, the number of sites will depend upon the species density on sites and the sensitivity required. Ideally, sample size should be large enough so there will be an 80 percent probability of being sure to detect a 20 percent difference when it exists. The size of the study site must be large enough to provide adequate samples. The size depends on the survey methods used, sensitivity required, and the density and range of the species of concern. For a paired plot design the number of sites required is a function of the average density of the species.

In general, the breeding density of the species of concern can be used to provide a rough estimate of the size of area needed to provide adequate samples. However, preliminary sampling most likely will be required to verify the estimates.

METHODS

Essentially, the methods used in a definitive study are a means to quantitate reproductive and mortality rates of animals on treatment and control areas. There are many texts and monographs available on methods of sampling to estimate these parameters (see Appendix A). Anyone not familiar with the theory and principles of the various techniques should review these references in depth. The objective of this section is to provide a general guide to the various methods that could be used in a definitive field study. In addition, these methods can be applicable to some screening studies.

The methods to be used in an individual field study will depend on the nature of the identified concerns. Some methods are useful for investigating several types of concerns; and most types of concerns can be studied by several methods. When the concern becomes more specific (*e.g.*, secondary hazards to raptors) or the use pattern and/or

habitat type is limited, the range of applicable methods tends to become more narrow.

Methods described below are divided into three categories: methods for assessing mortality and survival of adults and independent juveniles, methods for assessing reproduction and survival of dependent juveniles, and ancillary methods. The intent of this guidance document is to present methods that are likely to be useful in many situations, rather than an exhaustive list of all available methods. The Agency encourages the use of other methods when they are scientifically valid, and have a high probability of detecting an effect.

While it is absolutely essential to have a detailed investigational plan that describes the selected actions (with contingencies) for achieving the study objectives, investigators must remain flexible because anticipated problems always come up in long-term studies. Even with highly experienced and resourceful field biologists, the most carefully planned studies can be compromised due to the unpredictability of wild animals and natural events. When a natural disaster occurs early in the study, it may be wise to initiate the study again. If the event occurs after substantial data already have been collected (*e.g.*, early in the second year of a multiyear study), it may be more appropriate to extend the study an additional year or more to help provide for the additional needs. If the study is to be terminated, the report should describe thoroughly the nature of the event(s) and its (their) consequences if they (it) affect the study results.

MORTALITY AND SURVIVAL

It is very important to understand the autecology of the species being studied in order to select the most appropriate methods for investigating those species. In addition, the choice of particular methods must consider the applicability of the method based on the pesticide use pattern and study site characteristics.

Mark-Recapture

There are several mark-recapture methods available, each based on the same basic premise. A sample of animals is captured, marked, released, and another sample is collected where some of the animals are captured again. The characteristics of this identifiable sample then are used to estimate population parameters. Mark-recapture studies can provide information on:

- Size of the population;
- Age-specific fecundity rates;
- Age-specific mortality rates;
- Combined rates of birth and immigration; and
- Combined rates of death and emigration.

Seber (1982) reviewed the various mark-recapture methods and subsequent statistical analyses. Less detailed, but still very useful, reviews are provided by Caughley (1977) and Hanson (1967). Nichols and Pollock (1983) provide a valuable comparison of methods. Table 2 provides a brief summary of some of the various mark-recapture methods discussed in these references.

When considering the use of one of these mark-recapture models, one must carefully evaluate the applicability of the method to the circumstances under consideration. While in theory mark-recapture techniques should be an excellent method for evaluating effects of pesticides on wildlife populations, some mark-recapture analyses are not particularly robust; small deviations from their implicit assumptions can produce large errors in the results (Caughley, 1977). However, some of the more recent and sophisticated analytical

methods are robust and can deal with deviations from assumptions in closed populations (Otis *et al.*, 1978).

Mark-recapture methods are particularly useful for small mammals because these animals are seldom amenable to the visual and auditory observations necessary for using transect, territory mapping, or similar methods. However, mark-recapture also may be useful for birds provided a sufficient number of birds can be captured and marked. In some situations, birds may be "recaptured" with use of binoculars via visual observations of marked individuals.

Table 2.

Mark-Recapture Techniques.

Method	Applications / Requirements / Assumptions
Peterson Method (Lincoln Index)	Estimation of population size. Usually only two sampling periods. Closed Population.
Schumacher's Method	Estimation of population size. More than two sampling periods; marking continues throughout sampling. Closed population.
Bailey's Triple Catch	Estimate of birth rate and death rate in addition to population size. Requires data from two marking occasions and two recapturing occasions. Open population.
Jolly-Seber Method	Estimates mortality and recruitment in addition to population size. Requires more than two sampling periods and that each animal's history of recapture be known. Open population.

Animals must remain marked for the duration of the study. Typically, mammals are toe-clipped or ear-marked and birds are banded. Marking should not make the animals more susceptible to the effects of the pesticide (*e.g.*, anticoagulants with toe clipping). Dyes may be useful unless they are lost by wear or molting.

Territory Mapping Method

A common spatial census method is territory mapping, wherein the territories of individuals are mapped before and after treatment, on both treated and untreated plots. The method is usually applicable when birds are defending territories. It involves a series of census visits to the study sites during which birds located by sight or song are recorded on a map. The information from all the visits is plotted for each species. Birds exhibiting territorial behavior appear on the map as clusters of individual contacts. The clusters are used to estimate both the size and number of territories. The pre- and posttreatment censuses for treated sites are compared with the pre- and posttreatment censuses for control sites to determine changes in populations of territorial individuals that may be attributed to the pesticide (Edwards *et al.*, 1979). Further details of this method are given by the International Bird Census Committee (1970), and its application to evaluating impact caused by pesticides is reviewed by Edwards *et al.* (1979).

Problems with this method can occur. Under some circumstances, replacement from outside the area can be so rapid that territories are refilled before the census is completed. There usually is a floating population of silent, non-territorial birds who may quickly reoccupy empty territories (Stewart, 1951). The effects of replacement can be overcome for some species by capturing and marking the territorial individuals prior to treatment, so they can be distinguished from the floaters. Also, replacement may not be a problem when the study areas are in the center of a relatively large treated area.

Radio Telemetry

Radio telemetry can be an extremely useful technique to provide information on the effects of a pesticide application on nontarget species. As discussed for screening studies, radio telemetry can be used to monitor for mortality as well as to provide useful information on behavioral modification caused by the pesticide application. The points discussed previously (for screening studies) generally are applicable to definitive studies. However, for the definitive study, the number of radio-tagged animals needed depends upon the variation between sites and the sensitivity required. For example, with behavioral observation, intra- and inter-site variation will influence the number of radio-tagged animals required. In some instances, it might not be practical to radio-tag the number of animals required to provide a rigorously designed study. Under these conditions, the limitations should be specified, and the maximum number of animals that can be practically radio-tagged and monitored should be used.

Other Methods for Mortality and Survival

Other techniques for assessing density and diversity are discussed for screening studies; most of these, especially linetransect methods, are useful for definitive studies. Some methods, such as catch per unit effort or counts of animal signs, do not provide actual measures of density but may still be used to compare effects on treated and untreated plots.

REPRODUCTION AND SURVIVAL OF DEPENDENT YOUNG

Some of the techniques for assessing mortality and adult survival are also useful for assessing reproduction and survival of young. Some, but not all, mark-recapture methods can provide information on fecundity. Radio-tagging nestlings or suckling young of moderate and large size animals may be used to assess survival of dependent young. Radio telemetry and territory mapping are useful for locating dens or nests for further study. The following methods are more specific for assessing reproductive parameters.

Nest Monitoring

Nest monitoring is useful for evaluating the effect of pesticides on breeding birds. The typical procedure is to search the study site to find active nests and subsequently to check those nests to determine their fate. Information collected on each nest should include number of eggs laid, number hatched, number of young fledged, and if and when the nest was abandoned or destroyed, both before and after pesticide application. While all definitive studies should consider this technique, it also may be useful in screening studies.

This technique is relatively straightforward. However, it may not be practical if nests are scarce or otherwise hard to find. Because the breeding success of birds can be highly variable and can be quite low, it is sometimes difficult to obtain sufficient data on the success of the same species in enough sites to yield satisfactory results for statistical comparison with controls (Heinz *et al.*, 1979). In some cases, artificial nest structures can be constructed to increase nest densities. In a few situations where sufficient numbers are available, the technique may be applicable to mammal den monitoring.

Behavioral Observations

Behavioral observations associated with reproduction can be quite useful, especially for birds. Techniques are simple, but labor intensive. When used, such observations most likely would be combined with nest monitoring since both techniques require locating reproductive sites. Typically, the frequency and duration of behaviors will be compared for treated and untreated plots. Incubation, parental care (especially feeding for altricial birds), and following behavior (for precocial animals) are behaviors that are particularly

birds), and following behavior (for precocial animals) are behaviors that are particularly amenable to such study. Courtship, mating, and nest building are other behaviors that could be studied in some situations, but locating sufficient numbers of animals displaying these behaviors to permit quantitative analysis is difficult.

Age Structure of Populations

Comparisons of young-adult ratios of selected species between treated and untreated plots may indicate reproductive effects. The timing of the application and of breeding of selected species are critical. For assessing reproductive impairment or survival of dependent young, *per se*, the duration of this technique should be limited to single breeding-rearing periods, which may be repeatedly assessed. However, longer study periods that may even include several years can be used to assess the combination of reproductive success and age-specific mortality, even if the two cannot be separated.

Obviously, use of this method requires that the age of individual animals be determined. In some cases, it may be necessary only to distinguish among adults, sub-adults, and juveniles. In mammals, this may usually be accomplished by examining pelage, development of testes or mammaries, or tooth eruption or wear characteristics. In birds, plumage or characteristics of particular (species-dependent) feathers may be used. For carcasses or sacrificed animals, the ossification of bones or development of reproductive organs are useful. In other cases, particularly where comparisons are among populations in different years, it may be appropriate to distinguish age classes of adults. In mammals, tooth eruption, wear, or enamel layers, or eye lens weights are useful. It is more difficult to separate age classes of many adult birds, although overall plumage or feather characteristics can provide some indication. In some slow-maturing birds (e.g., gulls), plumage may be used to distinguish year classes of sub-adults. Additional details on aging birds and mammals are presented by Larson and Taber (1980).

ANCILLARY METHODS

At least some ancillary methods are essential in every field study. As used here, ancillary methods are generally of two types. Certain of these methods are important for determining the nature or existence of effects or for establishing causal relationships. Others of these methods do not address effects directly, but they provide important information for interpreting the results of the study.

Many of the methods for determining effects have been discussed for screening studies. Enzyme analysis, such as for cholinesterase inhibition, and observations of signs of toxicity can show that animals were exposed to or killed by a toxicant of a particular type. Where it is possible that animals may be exposed to other pesticides of the same type (*e.g.*, feeding in a nearby area treated with other pesticides), residue analysis in nontarget animals may be necessary to determine which specific pesticide caused the signs or alterations in enzymes. Even though carcass searches, *per se*, are not recommended for definitive studies, it is still essential to recover and analyze any carcasses found accidentally or obtained through radio-tracking. Residue and/or enzyme analysis of live animals collected will frequently be important.

Among the other ancillary methods, analysis of environmental residues is crucial and will probably be necessary in nearly every definitive field study. As discussed for screening studies, the most important environmental residues are those that occur on or in wildlife food sources, which may include insects, plant parts, or even other vertebrates, depending upon the species that are the primary focus of the investigation. The investigator should review the literature on food habits of the species being studied; often it will be appropriate to assess food habits on the specific study sites, particularly where the literature is not adequate to define food habits in the agricultural ecosystem under study. Such an assessment should include the availability of food sources and the amount

mobile animals that spend only part of the time in and adjacent to treated sites. The habitat should be thoroughly described to include both the morphology and species that are relevant to wildlife. Frequently, it will be important to locate and describe roosting, denning, or nesting sites for mobile wildlife that use treated sites part of the time.

INTERPRETATION OF RESULTS

Each field study is unique, although some elements may be common among many field studies. When a definitive field study is required, the requirement is based on one or more specific concerns that pertain to a specific chemical and one or several use patterns. Because of the substantial diversity in the types of problems to be assessed and the variety of available investigative methods, the key to understanding and interpreting a field study lies in the development of a sound protocol. All protocols will contain a description of the study sites, or the characteristics to be used in selecting sites within a given area, and the methods to be used in conducting the study. However, a well designed protocol will go beyond this descriptive approach in three ways.

First, the well-designed protocol will contain a restatement of the concerns to be addressed to ensure that there is an adequate understanding of the Agency's position. Then the investigator should review the literature and other available information that may bear upon the problem. It is possible that the literature may contain a valid answer to the questions raised by the Agency. Far more likely, the literature may orient the investigator to address the concerns in a particular way. An example is provided by Hegdal and Blaskiewicz (1984) who conducted a study to address the Agency's concerns for secondary toxicity to barn owls (specifically) from the use of an anticoagulant bait proposed for use on commensal rodents in and around agricultural buildings. A review of the literature by these investigators indicated to them that 1) laboratory studies suggested a legitimate potential for secondary poisoning to exposed raptors, but 2) the food habits of barn owls consist primarily of microtine rodents in most areas, suggesting a low potential for actual exposure. Consequently, they designed their study to focus on barn owl food habits and movements, and included an additive to the bait formulation that would permit an identification of whether or not the barn owls ate rodents that had fed on the bait. The study adequately demonstrated that actual exposure of barn owls was quite limited, and the proposed registration for this use was subsequently approved. By using the available literature on both the chemical and the particular species of concern, the investigators were able to narrow the study while still providing sufficient information for evaluation. However, it should be noted that this study was not adequate for evaluating the potential for secondary toxicity in the field to other predators that may have different food habits, or for other use patterns that may result in exposure to different predators or scavengers.

Second, the well designed protocol will contain reasons why particular methods are being used, including, at least qualitatively, the meaning that different results might have. For example, a protocol may include collection of residues in non-target animals, but it also should include a statement of purpose and meaning for such collection. Residues may be used to indicate potential exposure to nontarget organisms through analysis of their food, exposure in nontarget animals as a result of eating contaminated food, or that a particular pesticide was likely to be the cause of any observed effects. The interpretation of results is facilitated substantially by a statement of what is intended by using a particular technique. In the previously cited example from Hegdal and Blaskiewicz (1984), it was clearly stated that collection of owl pellets was to assess general food habits and that use of a fluorescing dye in the bait was for the purpose of ascertaining whether or not the owls fed on commensal rodents that specifically had fed on the bait. The interpretation of the data collected, once the purpose was stated, naturally led to the conclusion of no significant exposure to the barn owls.

Third, the well designed protocol will contain an experimental design that will indicate how the results can be assessed quantatively. The experimental design has been discussed in previous sections, but there are two facets that relate closely to the interpretation of results: the difference that can be detected between treated and untreated plots and the power (ability) of the design to detect this difference. Ideally, an experimental design with number of replicates based on an estimated coefficient of variation that closely approximates reality will allow the study to detect a stated concern level some prescribed number of times during the study time. Of course, the actual difference between treated and control units is measured during the field study, but the design can form an initial basis for interpretation when combined with the available information on the species of concern. As a result, the well designed protocol should include a section on interpretation.

Study methods for investigating acute mortality are more straightforward than for other kinds of effects. Nevertheless, there are sufficient differences in the use of the data to preclude a constant interpretation. The study may focus directly on the species of concern and may involve little or no extrapolation, depending on such factors as the type and the extent of use, the available toxicity data base, and home range of the species; or extrapolation to other populations, regions, or uses might be necessary. If the species of concern cannot be studied directly, it may be necessary to extrapolate between species, involving interspecies differences both in toxicological sensitivity and in ecological and population parameters.

The same kinds of considerations apply to reproductive impairment and chronic toxicity, even though different, and often more laborious and costly, investigative methods are involved. Where reproductive success is impaired, information on species-specific variation in reproductive ecology is necessary to understand how a particular degree of impairment may relate to effects among various species. Such reproductive considerations can include whether an avian species is a determinate or indeterminate layer, the number of nestings per season for different geographic areas in the use pattern, the length of the refractory period, as well as the specific effect which can range from destruction of reproductive organs to behavioral deficits such as nest abandonment. Considerations of reproductive ecology among different species of mammals include delayed fertilization or implantation, resorption of embryos or parental infanticide due to stress, number of young per breeding cycle, etc. All of these factors, and many others, are relevant to determining for different species the extent of effects that could result in population reductions or lack of ability to recover.

An analysis of whether or not a particular level of effect is going to affect wildlife populations is species-specific. For any species (or subspecies), the changes in population can be described very simplistically by the equation: rate of population increase (r) = birth rate - death rate, where r can be positive (population growth) or negative (population reduction). Where the concern is for specific populations of a species, then immigration and emigration are also important. These characteristics differ among species, and data will not always be available. The application of sound scientific judgment to the best available information will be the basis for interpreting the results of a study. It may be necessary to compare the results of the field study to laboratory data, especially where laboratory data are available on a variety of species and/or effects and the field study has focused on species other than those of direct concern. The use of extrapolation techniques will be necessary where endangered species are of concern or where other species cannot be studied directly.

Ideally, the Agency would like to be able to obtain a standardized result from a field study so that the result could be applied in a very consistent manner. As discussed in previous sections, the different effects and species of concern will vary and will require the development of specific protocols to address these factors. Although most of the various techniques have some degree of standardization, the field study may combine the individual techniques in a wide variety of ways to address specific concerns. A

standardized result might be attainable for the individual techniques, although that result would still have to be applied differently for various species, depending on their biology and ecological characteristics. However, determining a result for the whole field study that would unequivocally lead to a statement of the degree of risk, while obviously desirable, is not currently practical.

LITERATURE CITED

American Institute of Biological Sciences (AIBS). Aquatic Hazards of Pesticides Task Group. 1978. *Criteria and Rationale for Decision Making in Aquatic Hazard Evaluation.* Report to the Environmental Protection Agency. Contract No. 68-012457. August, 1978. 71 pp.

Balcomb, R. 1986. Songbird carcasses disappear rapidly from agricultural fields. The Auk 103(4): 817-820.

Bunyan, P.J., D.M. Jennings and A. Taylor. 1968a. Organophosphorus poisoning, some properties of avian esterases. J. Agr. Food Chem. 16: 326-331.

Bunyan, P.J., D.M. Jennings and A. Taylor. 1968b. Organophosphorus poisoning, diagnosis of poisoning in pheasants owing to a number of common pesticides. J. Agr. Food Chem. 16: 332-339.

Burnham, K.P., D.R. Anderson and J. L. Laake. 1980. Estimation of density from line transect sampling of biological populations. Wildl. Monogr., No. 72.

Caughley, G. 1977. *Analysis of Vertebrate Populations.* John Wiley and Sons, New York. 234 pp.

Cochran, W.W., 1980. Wildlife telemetry. pp. 509-520, in *Wildlife Management Techniques Manual.* S.D. Schemnitz, Ed. The Wildlife Society, Washington, D.C.

Corbett, J.R. 1974. *The Biochemical Mode of Action of Pesticides.* Academic Press, Inc., New York. 330 pp.

Davis, D.E. and R.L. Winstead. 1980. Estimating the numbers of wildlife populations. pp. 221-246, in *Wildlife Management Techniques Manual.* S.D. Schemnitz (Ed.), The Wildlife Society, Washington, D.C.

Eberhardt, L.L. 1985. Assessing the dynamics of wild populations. J. Wildl. Manage. 49(4): 997-1012.

Eberhardt, L.L. 1978. Appraising variability in populations studies. J. Wildl. Manage. 42(2): 207-237.

Eberhardt, L.L. 1976. Quantitative ecology and impact assessment. J. of Envir. Manage. 4: 27-70.

Edgington, E. 1980. *Randomization Test.* Dekker, Marcel Inc., New York, New York. 287 pp.

Edwards, P.J., S.M. Brown, M.R. Fletcher and P.I. Stanley. 1979. The use of a bird territory mapping method for detecting mortality following pesticide application. Agro-Ecosystems 5: 271-282.

Ellman, G.L., K.D. Courtney, V. Andres, Jr. and R.M. Featherstone. 1961. A new and rapid colorimetric determination of acetylcholinesterase activity. Biochem. Pharmacol. 7:88-95.

Environmental Protection Agency. 1982. *Pesticide Assessment Guidelines, Subdivision E - Hazard Evaluation: Wildlife and Aquatic Organisms.* Office of Pesticides and Toxic Substances, Washington, D.C. 87 pp.

Hanson, W.R. 1967. Estimating the density of an animal population. J. Res. Lepidoptera 6: 203-247.

Heath, R.G., J.W. Spann, E.F. Hill and J.F. Kreitzer. 1972. Comparative Dietary Toxicities of Pesticides to Birds. U.S. Fish and Wildlife Service. Special Scientific Report -- Wildlife No. 152., February. 57 pp.

LITERATURE CITED

Hegdal, P.L. and R.W. Blaskiewicz. 1984. Evaluation of the potential hazards to barn owls of Talon (brodifacoum bait) used to control rats and house mice. J. Environ. Toxicol. Chem. 3: 167-179.

Heinz, G.H., E.F. Hill, W.H. Stickel and L.F. Stickel. 1979. Environmental contaminant studies by the Patuxent Wildlife Research Center, pp. 8-35. in *Avian and Mammalian Wildlife Toxicology.* ASTM STP 693, E.E. Kenaga (Ed), American Society for Testing and Materials, Philadelphia.

Hill, E.F. and W.J. Fleming. 1982. Anticholinesterase poisoning of birds: field monitoring and diagnosis of acute poisoning. J. Environ. Toxicol. Chem. 1: 27-38.

Hill, E.F., R.G. Heath, J.W. Spann and J.D. Williams. 1975. Lethal Dietary Toxicities of Environmental Pollutants to Birds. U.S. Fish and Wildlife Service. Special Scientific Report -- Wildlife No. 191. 61 pp.

Hudson, R.H., R.K. Tucker and M.A. Haegele. 1984. *Handbook of Toxicity of Pesticides to Wildlife.* U.S. Department of the Interior. Fish and Wildlife Service. Resource Publication 153. 90 pp.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187-211.

Giles, R.H. Jr. 1978. *Wildlife Management.* W.H. Freeman and Company, San Francisco, CA. 416 pp.

International Bird Census Committee. 1970. Recommendations for an international standard for a mapping method in bird census work. Bull. Ecol. Res. Comm. 9: 49-52.

Kendall, M. and A. Stuart. 1977. *The Advanced Theory of Statistics* Volume 1, Distribution Theory, 4th Ed. MacMillan, New York. 472 pp.

Larson, J. S., and R. D. Taber. 1980. Criteria of sex and age. pp. 143-202, in *Wildlife Management Techniques Manual.* S. D. Schemnitz (Ed.), The Wildlife Society, Washington, D. C.

Leopold, A. 1933. *Game Management.* Charles Scribner's Sons, New York. 481 pp.

Ludke, J.L., E.L. Hill and M.P. Dieter. 1975. Cholinesterase (ChE) response and related mortality among birds fed ChE inhibitors. Arch. Environ. Contamin. Toxicol. 3: 1-21.

Moen, A.N. 1973. *Wildlife Ecology an Analytical Approach.* W.H. Freeman and Company, San Francisco. 458 pp.

Nichols, J.D. and K.H. Pollock. 1983. Estimation methodology in contemporary small mammal capture-recapture studies. J. Mammal. 64(2): 253-260.

O'Brien, R.D. 1967. *Insecticides: Action and Metabolism.* Academic Press, Inc., New York. 332 pp.

Otis, D.L., K.P. Burnham, G.C. White and D.R. Anderson. 1978. Statistical interference from capture data on closed animal populations. Wildl. Monogr. 62: 1-135.

Ripley, T.H. 1980. Planning wildlife management investigations and projects, pp. 1-6. in *Wildlife Management Techniques Manual.* S.D. Schemnitz (Ed.), The Wildlife Society, Washington, D.C.

Rosene, W. Jr. and D.W. Lay. 1963. Disappearance and visibility of quail remains. J. Wildl. Manage. 27: 139-142.

Scientific Advisory Panel (SAP). 1987. Final Scientific Advisory Panel subpanel's report on the January 7-8, 1987 meeting concerning terrestrial field studies. U.S. EPA, Washington, D.C.

Seber, G.A.F. 1982. *The Estimation of Animal Abundance and Related Parameters.* Macmillan Publishing Co., Inc., New York. 645 pp.

Shellenberger, T.E., B.J. Gough and L.A. Escuriex. 1970. The comparative toxicity of organophosphate pesticides in wildlife. pp. 205-210, in W.B. Diechmann, (Ed) *Pesticide Symposium.* Halos, Miami, FL.

Snedecor, G.W. and W.G. Cochran. 1967. *Statistical Methods.* Sixth Edition. The Iowa State University Press, Ames, IA. 593 pp.

Sokal, R.R. and F.J. Rohlf. 1969. Biometry. *The Principles and Practice of Statistics in Biological Research.* W.H. Freeman and Company, San Francisco. 776 pp.

Stewart, R.E. and J.W. Alrich. 1951. Removal and repopulation of breeding birds in a spruce-fir forest community. The Auk 68: 471-482.

Urban, D.J. and N.J. Cook. 1986. *Hazard Evaluation Divisions Standard Evaluation Procedure, Ecological Risk Assessment.* U.S. Environmental Protection Agency, Office of Pesticide Programs, Washington, D.C. 96 pp.

Walpole, R.E. and R.H. Myers. 1972. *Probability and Statistics for Engineers and Scientists.* The Macmillan Company, New York. 506 pp.

APPENDIX A

SELECTED REFERENCES

There are many publications available, ranging from journal articles to textbooks, that are pertinent to evaluating the impacts of pesticides to wildlife. The following list, although not exhaustive, is presented as a starting point for those not familiar with the subject area.

EXPERIMENTAL DESIGN AND STATISTICS

Cochran, W.G. 1977. *Sampling Techniques.* Third ed. John Wiley and Sons, Inc., New York. 611 pp.

Cochran, W.G. and G.M. Cox. 1957. *Experimental Design.* John Wiley and Sons, Inc., New York. 611 pp.

Green, R.H. 1979. *Sampling Design and Statistical Methods for Environmental Biologists.* John Wiley and Sons. New York 25711.

Johnson, D.H. 1979. Estimating nest success: The Mayfield method and an alternative. The Auk 96:651-661.

Otis, D.L., K.P. Burnham, G.C. White and D.R. Anderson. 1978. Statistical Inference from Capture Data on Closed Animal Populations. Wildl. Monogr. No. 62. The Wildlife Society, Washington, D.C. 135 pp.

Sokal, R.R. and F.J. Rohlf. 1981. *Biometry: The Principles and Practice of Statistics in Biological Research.* W.H. Freeman and Company, San Francisco. 859 pp.

Steel, R.G.D. and J.H. Torrie. 1960. *Principles and Procedures of Statistics.* McGraw-Hill, New York. 481 pp.

Walpole, R.E. and R.H. Myers. 1972. *Probability and Statistics for Engineers and Scientists.* The Macmillan Company, New York. pp. 506.

METHODS, GENERAL

Best, R.G. 1982. *Handbook of Remote Sensing in Fish and Wildlife Management.* Remote Sensing Institute, South Dakota State University, Brooking, South Dakota. 125 pp.

Blower, J.G., L.M. Cook and J.A. Bishop. 1981. *Estimating the Size of Animal Populations.* George Allen and Unwin Limited, Boston. 128 pp.

Caughley, G. 1977. *Analysis of Vertebrate Populations.* John Wiley and Sons, New York. 234 pp.

Copen, D.E. 1981. *The Use of Multivariate Statistics in Studies of Wildlife Habitat.* USDA Forest Service. General Technical Report RM-87. 249 pp. Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.

Davis, D.E. 1982. *Handbook of Census Methods for Terrestrial Vertebrates.* CRC Press, Inc.

Eberhardt, L.L. 1985. Assessing the dynamics of wild populations. J. Wildl. Manage. 49(4):997-1012.

Eberhardt, L.L. 1978. Appraising variability in population studies. J. Wildl. Manage. 42(2): 207-237.

Erickson, P.A. 1979. *Environmental Impact Assessment: Principles and Applications.* Academic Press, New York. 395 pp.

Gilmer, D.S., L.M. Cowardin, R.L. Duval, L.M. Mechlin, C.W. Shaiffer and V.B. Kuechle. Procedures for the Use of Aircraft in Wildlife Biotelemetry Studies. U.S.D.I. Fish and Wildlife Service Resource Publication 140. Washington, D.C. 19 pp.

Miller, H.W. and D.H. Johnson. 1978. Interpreting the results of nesting studies. J. Wildl. Manage. 42(3): 471-476.

Pettingill, O.S. 1970. *Ornithology in Laboratory and Field.* Burgess Publishing Company. Minneapolis. 524 pp.

Ralph, C.J. and J.M. Scott. 1980. Estimating Number of Terrestrial Birds. Proceedings of an International Symposium held at Asilomar, California, October 26-31, 1980. The Cooper Ornithological Society. 630 pp.

Schemnitz, S.D. 1980. *Wildlife Techniques Manual.* The Wildlife Society, Washington, D.C. 686 pp.

Southwood, T.R.E. 1978. *Ecological Methods: With Particular Reference to the Study of Insect Populations.* Chapman and Hall, New York. 524 pp.

Young, E. 1975. *The Capture and Care of Wild Animals.* Ralph Curtis Books, Hollywood, Florida. 224 pp.

MARK-RECAPTURE

Brownie, C., D.R. Anderson, K.P. Burnham and D.S. Robson. 1985. *Statistical Inference from Band Recovery Data-A Handbook.* USDI, Fish and Wildlife Service Resource Publication No. 156. Washington, D.C. 305 pp.

Hanson, W.R. 1967. Estimating the density of an animal population. J. Res. Lepidoptera 6:203-247.

Nichols, J.D. and K.H. Pollock. 1983. Estimation methodology in contemporary small mammal capture-recapture studies. J. Mamm. 64(2):253-260.

Otis, D.L., K.P. Burnham, G.C. White and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildl. Monogr. 62: 1-135.

Seber, G.A.F. 1982. *The Estimation of Animal Abundance and Related Parameters.* Macmillan Publishing Co., Inc., New York. 654 pp.

White, G.C., D.R. Anderson, K.P. Burnham and D.L. Otis. 1982. *Capture-Recapture and Removal Methods for Sampling Closed Populations.* Los Alamos National Laboratory, Publ. LA- 8787-1-NERP, Los Alamos, NM. 235 pp.

LINE TRANSECT

Best, L.B. 1981. Seasonal changes in detection of individual bird species. Studies in Avian Biology. 6:252-261.

Burnham, K.P., D.R. Anderson and J.L. Laake. 1985. Efficiency and bias in strip and line transect sampling. J. Wildl. Manage. 49(4):1012-1018.

Burnham, K.P., D.R. Anderson and J.L. Laake. 1980. Estimation of Density from Line Transect Sampling of Biological Populations. Wildl. Monogr. No. 72.

Crain, B.R., K.P. Burnham, D.R. Anderson and J.L. Laake. 1978. *A Fourier Series Estimator of Population Density for Line Transect Sampling*. Utah State University Press. 25 pp.

Eberhardt, L.L. 1978. Transect methods for population studies. J. Wildl. Manage. 42(1): 1-31.

Mikol, S.A. 1980. Field Guidelines for Using Transects to Sample Non-Game Bird Populations. U.S.D.I, Fish and Wildlife Service Pub. 80/58. Washington, D.C. 26 pp.

TERRITORY MAPPING

Best, L.B. 1975. Interpretational errors in the "mapping method" as a census technique. The Auk. 92(3):452-460.

Edwards, P.J., S.M. Brown, M.R. Fletcher and P.I. Stanley. 1979. The use of a bird territory mapping method for detecting mortality following pesticide application. Agro-Ecosystems. 5:271-282.

International Bird Census Committee. 1970. Recommendations for an international standard for a mapping method in bird census work. Bull. Ecol. Res. Comm., 9:49-52.

ANTICHOLINESTERASE

Grue, C.E., G.V.N. Powell and C.H. Corsuch. 1982. Assessing effects of organophosphates on songbirds: comparison of a captive and a free-living population. J. Wildl. Manage. 46(3):766-768.

Hill, E.F. and W.J. Fleming. 1982. Anticholinesterase poisoning of birds: field monitoring and diagnosis of acute poisoning. J. Environ. Toxicol. Chem. 1:27-38.

APPENDIX B

SUGGESTED COMPONENTS OF A FIELD STUDY PROTOCOL

for Submittal to EEB for Review.
Adapted from Ripley (1980)

I. Title

II. Problem Definition

A. A review and summary of the available information on the pesticide in relation to nontarget hazard, including use information.

B. A precise statement of the goals and purpose of the study(ies) (objective(s)).

C. A brief statement of the problem and the context in which it exists, specifying the limits of the proposed work (Scope).

D. Precise statements of the major hypotheses to be tested.

III. Methods and Materials

A. A brief discussion of various methods and procedures that have been or could be used to evaluate the problem. This discussion should identify the strengths and weaknesses of each method or procedure discussed.

B. Descriptions

1. Identify the study area(s) selected and their general suitability for achieving the objectives of the study or what criteria will be used to select study areas.

2. Identify the species present or expected to be present on the study area(s), discussing characteristics pertinent to the problem being evaluated.

3. State the research procedures, designs and sampling plans to be used.

a. Specify the kind and amount of data needed and to be sought.

b. Describe in detail how all data are to be obtained, including details of application, instrumentation, equipment, sampling procedures, etc.

4. Describe how the data are to be treated, including specifying what statistics are to be calculated, what models will be used, what tests of data will be used, etc.

5. Describe in detail the methods to be used to check the sensitivity and accuracy of the procedures used.

6. Describe Quality Assurance procedures for application, instrumentation, equipment and records.

7. Briefly describe the resources (people, facilities, etc.) to be applied to the study.

APPENDIX C

CARCASS SEARCHES

DESIGN

In designing carcass searches, the following factors need to be known or determined:

- Density of the species that are likely to be exposed. For example, granular products are most likely to result in exposure to ground-feeding animals; therefore, birds such as warblers or swallows should not be included in density counts for such products;
- Probability of finding (a) dead animal(s) if any are killed. This is dependent on the probability of a carcass remaining on the study site (i.e., not being removed by scavengers) and the probability of detecting a carcass if it remains on the study site (search efficiency);
- Size of the search area; and
- Number of carcasses found.

These factors can be combined in the following formula:

$$N = D\,R\,E\,A\,P$$

where

N = number of carcasses found

D = density in animals/acre

R = proportion of carcasses remaining (nonremoval)

E = search efficiency

A = acres searched

P = proportion of population killed

Carcass searches should be used only when there is a reasonable potential to detect mortality. If such mortality does occur, the carcass search should be able to detect it and therefore, carcasses should be found. It is recommended that carcass searches be designed so that at least two carcasses (N = 2) will be found if there is appreciable mortality. In general, preliminary sampling would be required to determine these factors. However, information from other field studies can be used in the planning stages to determine if carcass searching would be appropriate for use under anticipated conditions and to assist in developing the study design.

The sensitivity of the carcass search approach is equivalent to the percent detectable kill of the population. To determine the sensitivity, the formula is adjusted:

$$P = \frac{N}{D\,R\,E\,A}$$

Since P is a proportion:

$$\text{percent detectable kill} = P \times 100 = \frac{N}{D\,R\,E\,A} \times 100$$

Obviously, if any of D, R, E or A are zero, the equation cannot be solved and the carcass search is not applicable (i.e., no density of birds, no acres searched, no carcasses

remaining, no remaining carcasses found). However, other combinations of D, R, E and A, such as low density and small acreage or low efficiency and high scavenger removal, can result in a small denominator meaning that mortality can be detected only when a high percentage of the population is killed. For example, in 5 acre fields with only 2 birds/acre and R and E estimated at a moderate 0.5, only an 80 percent or greater kill could be detected. In such situations, it is necessary to increase one of the parameters to achieve a stated level of detectability or else to use methods other than carcass searching. The same equation can be used to estimate the minimum search area to detect a given mortality level (P) by solving for A.

SEARCH PROCEDURE

In general, depending on the sensitivity of the search method relative to the habitat involved, corridors or plots should be selected. These areas should be searched systematically by walking predetermined routes until the area has all been covered. Due to the concentration required to find dead animals, other activities that could distract the searchers' attention should be avoided during carcass searching. In homogeneous situations, investigators should randomly select search areas. However, in most studies it is advisable to stratify the sampling, concentrating efforts in areas frequented by wildlife species such as woods edges, ditch banks, field borders, fencerows and other habitats where wildlife concentrate.

DURATION

Searches should begin on the day of application and continue on a daily basis for as long as mortalities or other evidence of intoxication occur. In general, a week or two following application should be adequate. However, the length of time searches are continued should be related to how long lethal concentrations are expected to be present. Normally, the same areas should be searched each day.

ESTIMATING EFFICIENCY OF CARCASS SEARCH

Efficiency trials should be conducted periodically (minimum 3 times per study site) during the study to determine the proportion of carcasses that are detected. Just prior to the initiation of a scheduled search, carcasses of species representative of species found in the area should be variously placed within the search area. If the study site includes edge habitat, carcasses should be placed in the edges as well as in the fields. In general, carcasses should be placed where animals would be most likely to die, depending on the nature of the chemical. Searchers should not be aware that simulated mortalities have been placed; however, they should be aware that these trials will occur during any scheduled search.

The number of carcasses placed should be approximately equal to 20 percent of the estimated density of species on the search area. All placed carcasses should be marked to distinguish them from actual kills. The location of placed carcasses should be mapped so those not found can be easily recovered following completion of that day's search activities, since unrecovered carcasses could bias study results. For example, if a scavenger were to carry off a simulated kill and consume it at another location on the study site, the remains could be erroneously classified as pesticide-related if found. One potential way to avoid this problem would be to dip carcasses in a nontoxic substance that fluoresces under ultraviolet light so that the remains could be identified.

ESTIMATING CARCASS REMOVAL RATE

Carcass removal should be monitored to determine local variability in scavenger activity. The density of both carcasses and scavengers can influence the rate of removal. Under some conditions, large numbers of carcasses may attract scavengers. In other situations a large number of kills may dilute removal rate due to limited number of scavengers. Where it can be adequately documented that removal of carcasses occurs almost exclusively either at night or during the day, the timing of carcass searches may be adjusted to minimize the effects of removal.

Carcasses planted in monitoring trials should simulate mortalities actually occurring from the pesticide. In most cases, small to moderate sized species such as starlings or blackbirds, or laboratory bobwhite or Japanese quail chicks may be used. Carcasses should be variously placed within the general study areas and monitored daily for at least 5 days or until 90 percent have been removed. Ideally, the number used should approximate densities resulting from effects of the pesticide under study; however, in most instances, this will not be known. Therefore, a density of approximately 20 percent of the population of nontarget species on the area is recommended.

Timing of carcass removal trials should be such that they do not affect scavenger removal of pesticide-killed birds or the feather-spots of the removed carcass could be erroneously classified as a pesticide kill. Location of placed birds should be recorded on maps and may be marked in the field with small stakes or by other inconspicuous means, preferably at a fixed distance and direction from the carcass.

APPENDIX D

EXAMPLES OF METHODS AVAILABLE FOR INVESTIGATING PARTICULAR, IDENTIFIED EFFECTS

Every field study must address specific concerns for a specific chemical in a specific use pattern. Just as each chemical differs at least slightly from other chemicals, each field study is likely to differ at least slightly from other field studies. It is impossible to provide thorough directions or methodology that will apply to all pesticides, including those yet to be developed. However, some of the kinds of information required in field studies can be related to the types of concerns or effects that have been identified from lower tier testing or other information. Table D-1 provides a general list of methods that are most likely to be appropriate for addressing typical concerns. Following the table is a discussion of methods for two chemicals. For some pesticides and use patterns, there may be more than one kind of identified concern; in such situations, the field study methodology should be able to address all of the identified concerns.

One critical aspect of field studies is not considered in the following discussions. The kinds of techniques used for investigating effects are less important than the validity of and within the methods for elucidating effects, including lower tier data. The use of every conceivable technique is ineffectual, not to mention very costly, if sites are inappropriate, application rates are low, sampling design will allow only a low probability of detecting an effect, exposure (or lack thereof) is not documented, etc.

A second critical aspect is subjective. How accurately can the investigator predict the results of the field study or its various aspects? Using acute mortality as an example, if the investigator is nearly certain that field mortality will not occur (for whatever reasons), then a screening study would be not only appropriate, but also cost-effective. Conversely, if the investigator believes that there is a likelihood that mortality will occur above concern levels, then a screening study may be a waste of time and money, except that it might have utility as a baseline study for the forthcoming definitive study. Similarly, the requirement for, or nature of, a field study may depend on the environmental concentrations, especially in or on wildlife food resources. Although residue estimation techniques have frequently been shown to be reasonably accurate, there are some situations where estimations are far from measured residues. If the investigator is genuinely confident that actual residues are far less than estimated, to the extent that a requirement would be removed, then actual residue data should be obtained to provide a more cost-effective measure of likely effects. But, as in the previous example, there is little point in obtaining such data prior to a field study if the investigator predicts that actual residues will be similar to estimated residues.

EXAMPLE 1

A cholinesterase inhibiting ("irreversible") compound with the potential for causing mortality quickly after ingestion of environmentally relevant amounts. Avian reproductive tests show reduced productivity of young apparently as a result of parental toxicity, but no evidence of impairment of other reproductive processes.

Discussion- If the reproductive effect levels are above the environmental levels, this investigation would focus on acute mortality. Unless there are several documented kills from normal use, a screening study would be the likely approach. Such a study would be oriented towards both birds and mammals unless acute toxicity data indicate one of

Table D-1.

Priority (1=highest priority) of field study methods for addressing different types of concerns or effects.

Method	Lethality[1]		Sublethal[2]		
	Fast	Chronic or Delayed	Acute	Chronic	Resource Loss[3] (Food or Cover)
Environmental Residues[4]	1	1	1	1	1
Nontarget Residues[4]	2	1	1	1	
Density and Diversity[5]	1	2	2	3	1
Enzyme Analysis	2	1	1	1	
Behavioral Observations	2	1	1	1	2
Nest and Nest Box Monitoring	3	2	1	1	1
Carcass Searching	1	3			
Radio Telemetry[6]	2	1	2	2	2
Mark-recapture[6]	3	1	2	2	2
Adult : Young Ratios	3	2	2	2	2
Gross Histopathology			2	2	
Resource Survey					1

Footnotes

[1] Lethal responses may be fast (e.g., onset of mortality is less than 1 hour in laboratory studies) or delayed (onset 12 hrs. or more) or intermediate. The same techniques are useful, in general, for any type of lethal response except that carcass searching is most useful with fast-acting compounds (< 1 hour onset) and decreases in utility with slower acting compounds. Carcass searches are of questionable validity (with exception for unique situations), when the onset of mortality is greater than 12 hours. Conversely, although radio telemetry and mark-recapture techniques may be useful for a fast-acting compound, they become increasingly useful when the time-to-toxicity is intermediate and essential for delayed or chronic lethal responses.

[2] Defined as a response to a single or repeated application of a pesticide that reduces the fitness of a nontarget organism to survive, reproduce or rear young. Acute effects may render an organism more susceptible to predation, cause nest abandonment or spontaneous abortion, impair the ability of adults to feed their young, etc. Chronic effects may be similar in nature or involve more subtle effects on reproductive potential, but they are manifested more slowly.

[3] Loss of resources, such as nontarget organisms that provide food or cover for wildlife, typically does not provide a basis for conducting field studies. However, in some circumstances, particularly where large contiguous acreage is treated, effects on food or cover may be pronounced and may warrant a field study.

[4] As used in this table, environmental residues mean residues in soil, water, wildlife food resources, etc. Nontarget residues are those residues found in wildlife (either found dead or collected live) that may indicate the toxic agent causing effects. In some cases, such as where there is a concern for both primary and secondary toxicity, the same animal (e.g., a mouse or sparrow) may be part of the nontarget residue collection for determining cause of effects and also part of the environmental residue collection as a food source for a predator. If existing chemical fate data are inadequate to assess changes in environmental residues, sampling at several intervals may be very useful.

[5] All field studies must provide some description of the species, numbers and nature of utilization of nontarget wildlife associated with the study site(s). Such a description is essential during the site selection process to ensure that the study can provide useful information. This entry on the table refers to density and diversity estimates that are made during the actual study and may be fairly general (e.g., for the purpose of determining the size of the carcass search area) or fairly detailed (e.g., when the estimates are used to compare changes in populations pre- and post-application or between treated areas and controls).

[6] Typically, as the usefulness of carcass searches for assessing lethality decreases with longer time-to-toxicity for a pesticide, the utility of radio telemetry and/or mark-recapture methods increases. Mark-recapture and radio telemetry normally will not both be used in the same study. Mark-recapture is most useful with moderately common species with low mobility (e.g., small mammals), whereas radio telemetry is most useful for less common or more mobile species (e.g., raptors especially, but also birds in general and moderate size mammals).

these taxa to be much less sensitive than the other. After sites are selected where there is an adequate abundance and diversity of wildlife, appropriate techniques would include:

- Carcass searching for birds and mammals, cover permitting. If carcass searching is not feasible for mammals or birds, mark-recapture or radio telemetry are useful alternatives.
- Collection of environmental residues.
- Cholinesterase assays to assess cause of mortality, supplemented by residue analysis if other cholinesterase inhibitors are used in the area.
- Density and diversity estimates for use in calculating search area size and probability of detecting dead wildlife.

If there is a question about the environmental levels, relative to reproductive effect levels, the collection of environmental residues during a screening study may permit an assessment of the potential reproductive effects in the field. If reproductive effect levels are lower than environmental concentrations as determined either by data collection during a screening study or through acceptable estimation techniques, then a definitive field study would be appropriate for assessing such effects. However, different approaches would be used depending, first, upon whether or not a screening study had been conducted and, second, the results of the screening study.

EXAMPLE 2

A valid screening study showed no acute mortality. Concerns would be focused on potential reproductive impairment in the field. Appropriate techniques should include:

- Additional environmental residues.
- Nontarget residues in live-collected wildlife.
- Cholinesterase assay in collected wildlife.
- Behavioral observations, particularly related to reproductive and nesting/parental behavior.
- Monitoring of nest/dens or artificial structures to evaluate productivity relative to control sites.
- For mammals, evaluation of young-adult ratios relative to control sites and/or pretreatment.
- Depending on the use pattern and nature of the test plots, radio telemetry and/or mark-recapture techniques may be useful alternatives.

A valid screening study indicated greater than the concern level for mortality occurred over a stated percent/sites. Concerns and techniques would be as above plus additional techniques should be employed to determine the extent and importance of acute mortality.

EXAMPLE 3

- Quantitative density and diversity methods for treated (including nearby habitat) and control sites.
- Mark-recapture methods may be particularly useful for mammals.

- Radio telemetry has some disadvantages (primarily the number of organisms required) for quantitative acute effects, but could be useful for this purpose if already being used to investigate productivity parameters.

EXAMPLE 4

No screening study has been done. Environmental residues, either from actual data or from acceptable estimation techniques, exceed both acute and reproductive effect levels. If the actual existence of acute effects and the estimation of environmental residues are questionable, a screening study may be useful, but, unless both residues are lower than reproductive effect levels and no mortality is found, the screening study would have to be followed by a definitive study. Unless the investigator was quite confident that a screening study would be "clean" on both counts, it would be quicker and more cost-effective to proceed directly to a definitive study. Since a definitive study for assessing reproductive effects is nearly always a multi-year study, the assessment of acute effects in the first year could be of the screening type. If effects above concern levels are found, a more thorough assessment of acute effects may be made in the second and/or subsequent years. Appropriate techniques for both acute and reproductive concerns have been listed above. However, because both concerns would be investigated at the same time, the investigator should consider carefully how these techniques can be combined in the most efficient manner.

EXAMPLE 5

An anticoagulant rodenticide causes mortality after a delay of one to several days (regardless of whether due to one or several feedings). In addition to nontarget mortality from primary exposure, there is a concern for secondary toxicity to predators or scavengers feeding on either dead or live target rodents.

Discussion- Anticoagulants frequently have quite different toxicity to different taxa of wildlife. Concerns for secondary toxicity may be based on reasonable scenarios or on known incidents of secondary poisoning and the concerns may be for a broad or narrow array of secondary consumers. If concerns are for one taxon (e.g., buteonid raptors or wild canids) and are based upon potential, rather than known effects, laboratory studies on secondary toxicity would be strongly recommended and should provide accurate information on residues in primary consumers as well as toxicity to the secondary consumer. Assuming that this laboratory study supports the potential for field effects and provides dose-response information (or a NOEL), the residues in primary consumers (equals secondary exposure levels) are important in interpreting any field results. It is essential that any secondary toxicity field study include considerations of food habits of the secondary consumer.

With regard to primary nontarget toxicity, it can be assumed that a rodenticide will kill nontarget rodents and probably other nontarget mammals that ingest the toxicant. Birds, as primary consumers, may or may not be particularly sensitive, but if extended laboratory studies indicate they are, they should be included in the field study design. Except where the sensitivity of birds is equivocal, with respect to exposure, there is little point in screening studies for a vertebrate toxicant; this example compound is designed to kill rodents. Appropriate techniques for this example compound include:

- Behavioral observations, especially regarding food habits of consumers.

- Residues in target and nontarget primary consumers and possibly secondary consumer.

- Mark-recapture for small nontarget mammals.

- Radio telemetry for secondary consumers, larger nontarget mammals and birds; effects on birds could also be studied through density and diversity (i.e., census) methods or mark-recapture.

- Insofar as possible, target and nontarget carcasses should be collected for analysis, but systematic carcass searches are of little use when mortality is delayed.

APPENDIX E

TERRESTRIAL FIELD STUDIES, WHEN ARE THEY REQUIRED?

GENERAL

The Agency utilizes a tiered system of ecological effects (usually toxicity) testing to determine the potential risks of proposed pesticide uses to nontarget aquatic and terrestrial organisms. These tests are outlined in various subdivisions of the Guidelines with Subdivision E addressing the pesticide's effects to birds, mammals and aquatic vertebrates and invertebrates; Subdivision J addressing nontarget plant effects; and Subdivision L addressing nontarget insect effects. However, the terrestrial toxicity or adverse effects, data usually available for risk assessments are as follows:

Tier 1

Mammalian Toxicity Data

Submitted in support of (human) toxicology data requirements (*e.g.*, rat acute oral LD_{50}; acute dermal toxicity; 90-day feeding studies -- rodent and nonrodent §'s 81-1 through -7; 82-1 through -5; 83-1 through -4; 84-2 through -4; and §'s 85-1, 85-2 and 86-1).

Avian Toxicity Data

Avian acute oral LD_{50} (upland gamebird or waterfowl species) (§71-1); Avian dietary LC_{50} (upland gamebird) (§71-2); and Avian dietary LC_{50} (waterfowl species) (§71-2).

Tier 2

Wild Mammal Toxicity Data

Generally, a dietary LC_{50} or acute oral LD_{50} study with a non-endangered representative species that is likely to be exposed (§71-3).

Avian Reproductive Studies

Studies using upland gamebird and waterfowl species (§71-4).

Special Studies

Studies with specified avian or mammalian species such as nontarget mammalian reproduction studies, avian acute dermal LD_{50}, avian cholinesterase test, avian or mammalian secondary toxicity (§70-1).

Tiers 3 and 4

Field Tests

Simulated and/or actual field testing with avian and/or mammalian species (§71-5).

Test Species

The typical mammalian and avian indicator species used in the toxicity tests above are the domestic rat, bobwhite quail, ring-necked pheasant and mallard duck. Other organisms such as cottontail rabbits, voles and songbirds are sometimes used on a case-by-case basis to address specific risks. Generally, those organisms, representative of areas where pesticide applications are likely to occur, are utilized (excepting endangered species).

ECOLOGICAL/TERRESTRIAL RISK ASSESSMENT

In order to determine when terrestrial field studies are required to support a pesticide use proposed for registration, the Agency must perform an ecological risk assessment. This process is composed of two major areas: an aquatic risk assessment and a terrestrial risk assessment. The Agency also assesses the risks to nontarget plants and to nontarget invertebrates (primarily, to beneficial insects such as honey bees). However, since the aquatic and, especially, the terrestrial assessments are the major determinants of when terrestrial field studies are required, they will be discussed in detail here.

Components of Ecological Risk

The components of both the terrestrial and aquatic risk assessments can be presented as follows:

Toxicological Hazards	X	Environmental Exposure	=	Ecological Risks

or

Effects (Toxicity) Data	X	Exposure Data	=	Estimates of Ecological Risks

Table E-1 breaks this relationship down further to show the data and/or information utilized.

As required by FIFRA, when the Agency performs an ecological risk assessment, it performs the terrestrial and aquatic risk segments together. The terrestrial assessment has the greatest impact on determining when terrestrial field studies are required, but the aquatic segment is an important element that could show a need for such studies. For example, if a pesticide use provided for adverse impacts on aquatic food sources and the Agency estimated that such impacts could adversely affect nontarget terrestrial organisms, then a terrestrial field study might be required.

A similar discussion, relative to adverse effects on wildlife habitat and terrestrial food items (e.g., invertebrates such as earthworms, insects, slugs), can be presented. However, although the Agency tries to address these areas in its ecological risk assessment, the EPA focuses on the acute, subacute and/or chronic risks to mammals, birds and aquatic vertebrates and invertebrates via ingestion, dermal exposure, inhalation and/or aquatic exposure. It does not usually address adverse effects via loss of habitat or from loss of terrestrial food items unless endangered species are involved or catastrophic losses appear

Table E-1.

The relationship between the components of ecological risk can be broken down to show the type information utilized:

Toxicological Hazard Data	X	Exposure Data	=	Ecological Risks
- Laboratory eco-toxicity data (e.g., LD_{50}, LC_{50} and NELs)		- Physical and chemical properties		- Integration of data into Agency statement of risk for both endangered and non-endangered species
- Human Toxicity data (e.g., NELs, Chronic effects)		- Chemical fate and transport data		
- Field data (sometimes available) or, small pen avian and mammalian species		- Nontarget organism and habitat information for both endangered and non-endangered species		
- Pesticide incidents data (e.g., avian field kills)		- Pesticide use information - Pesticide residues (estimated and/or actual)		

likely, based on Agency estimates or a body of data that shows that such losses are possible.

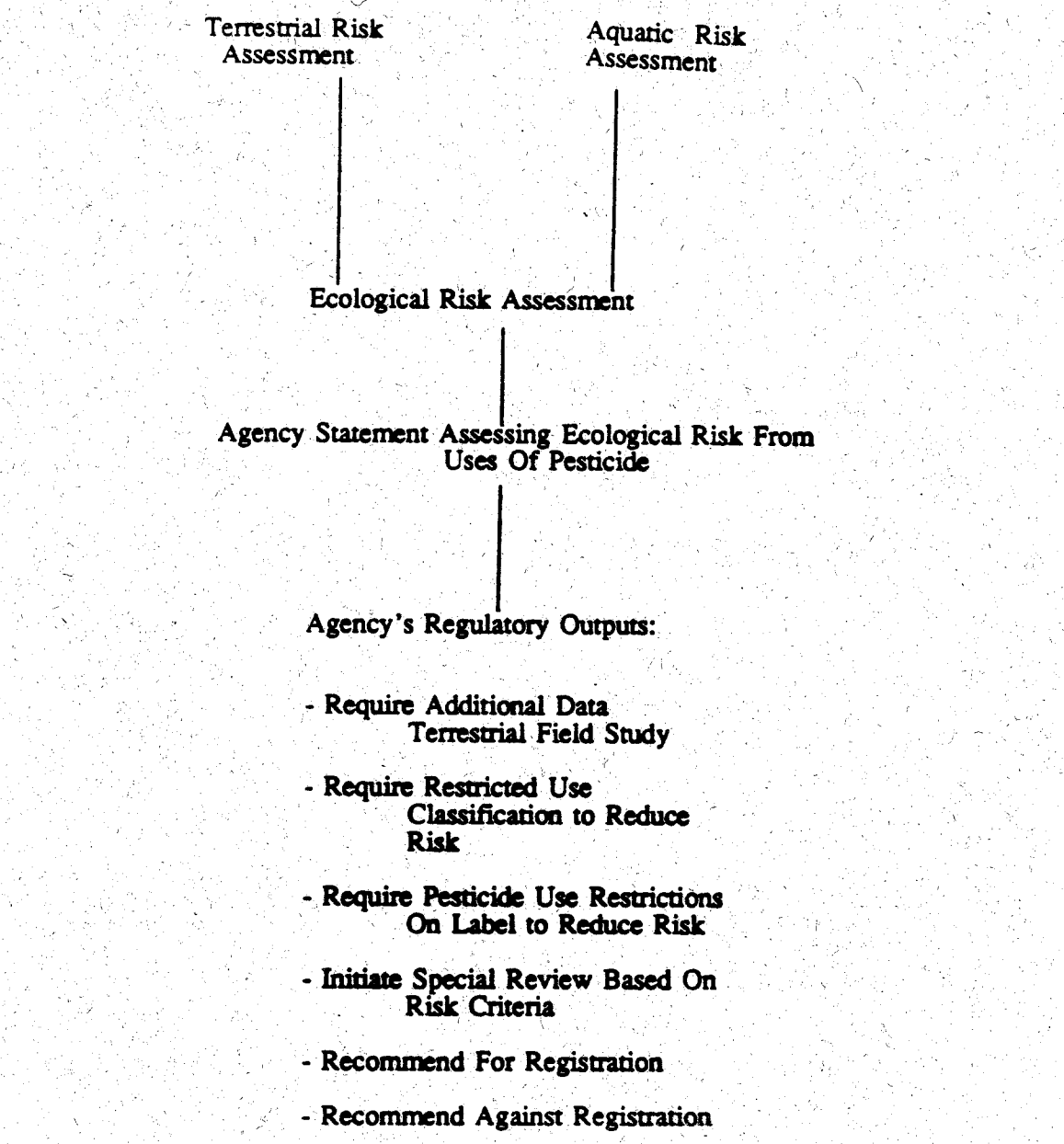
The Agency does this because the largest and often times best, effects data base available is the toxicity/effects data for mammals, birds and aquatic vertebrates and invertebrates. As the state of the art improves, however, EPA will perform more ecosystem-level risk assessments utilizing the effects data for all ecosystem components.

The process that usually generates the requirement for a terrestrial field study is the Agency's ecological risk assessment, especially its two major components the aquatic assessment and the terrestrial assessment. Figure E-1 is a schematic presentation of how the Agency moves from this assessment process to the field study requirement.

The terrestrial risk segment of the ecological risk assessment is usually the major area that "triggers" the requirement for a terrestrial field study. The terrestrial risk assessment process examines the potential risks of pesticide uses to non-human, nontarget terrestrial organisms; primarily, to nontarget mammals and avian species. Reptiles and amphibians are not necessarily ignored, but it is assumed that when birds and mammals are "protected," via Agency risk procedures and criteria, that some "protection" is afforded reptiles and amphibians. Further, as the state of the art of toxicity testing develops, other organisms, such as reptiles and amphibians, can be considered more accurately in the risk assessment process (Urban and Cook, 1986).

Figure E-1.

A Schematic Presentation of How EPA Moves from the Risk Assessment Process to Field Study Requirements.



EXPOSURE DATA

Under the exposure portion of the relationship, Toxicological Hazards X Environmental Exposure = Ecological Risks, the Agency examines five areas:

Physical/Chemical Properties

The EPA requires that this information be submitted for all pesticides and, generally, the information of most value in the ecological risk process includes: color, physical state, melting point, vapor pressure, density or specific gravity, solubility, dissociation constant, octanol/water partition coefficient, pH, molecular weight and chemical structure (Urban and Cook, 1986). These data give the Agency an indication whether the pesticide is highly soluble, volatilizes readily, is related in chemical structure to other pesticide compounds, etc.

Chemical Fate/Transport Data

The data submitted to the Agency and typically utilized in the risk process includes: hydrolysis, photodegradation in water or on soil, metabolism studies, leaching potential, field dissipation (residue decline curves, metabolites) and bioaccumulation. Using these data the EPA estimates which potential exposures are likely: acute, subacute, chronic (reproductive) and/or secondary or, possibly, tertiary because of build-up in the food chain (Urban and Cook, 1986).

Pesticide Use Information

Generally, much of the pesticide use information is submitted by the pesticide applicant. However, the Agency also examines public literature that may provide pertinent data (e.g., the U.S. Department of Agriculture's Agricultural Statistics handbook) on the proposed use. Information that is factored into the terrestrial risk assessment process includes: the type of formulation (granular, wettable powder, flowable, microencapsulated), type of application equipment (helicopter, plane, ground), crop acreages to be treated, amount of pesticide to be applied (amount per acre, low volume, high volume, ultralow volume (ULV)), timing of application (time of day, time of year), number of applications per season, intervals between applications, use site(s), target pest(s), inerts in the formulation and diluents, surfactants, adjuvants or stickers used, if any.

Nontarget Organism/Habitat Information

The Agency, primarily through the use of its staff expertise, the public literature and contacts with the U.S. Fish and Wildlife Service (USFWS), the Office of Endangered Species (OES), State fish and game agencies, academicians and other experts in the field, determines the nontarget avian and mammalian species, including reptiles and amphibians, when possible, are likely to be exposed. Both non-endangered and endangered organisms are considered including: what species/habitats are exposed; what life stages are exposed; for how long does exposure occur; whether exposure is acute, intermittent or chronic; and what food sources may be contaminated. For federally listed endangered species the OES or the National Marine Fisheries Service (NMFS) is contacted via informal and formal consultation procedures. Documents, information and data are forwarded to OES (or

NMFS) during the formal consultation and the OES responds with a formal Biological Opinion identifying those endangered species likely to be affected by the pesticide use pattern(s). This Opinion is reviewed by the EPA and recommendations, based on the opinion and relative to the pesticide use(s), are developed by the Agency. It should be noted that every attempt is made by the EPA to protect federally listed (and, when possible, State listed) endangered species.

Actual Pesticide Residues

The Agency has access to up to four data bases for actual pesticide residues. These are:

Chemical Fate/Transport Data

Pesticide residues in the form of half-life estimates, actual measured residues and residue decline curves are generally available in this data package. Such data are submitted by the pesticide registrant and are reviewed and validated by the Agency. Said data are heavily utilized in the ecological risk assessment process.

Residue Tolerance Data

Actual residue data are required by the Agency for pesticides used on crops that may be consumed by humans and/or domestic animals (such as cattle) or on crops that may be processed into human and/or cattle food or feed items. Also, residues in/on fish and shellfish are required to support pesticides used in aquatic sites. These residue data, however, are usually of limited utility to the ecological risk assessment process because such data generally consist of residues determined at the time of crop harvest and for crop items consumed by humans, but not necessarily by nontarget wildlife. These data are developed for use in the human risk assessment process; however, when possible, the Agency utilizes said information in its ecological risk assessment.

Residues In/On Wildlife Food Items

Occasionally, but rarely, data for pesticide residues in/on wildlife food items such as insects, other invertebrates, seeds, pods, forage or nuts are submitted by the registrant to the Agency prior to the determination that a terrestrial field study is required. However, the Agency normally does not request such data until it determines that a terrestrial field study is needed to assess the risks.

Public Literature

Whenever possible, the Agency utilizes actual pesticide residue data found in the literature. Often, however, such data are lacking, particularly residues in/on pertinent wildlife food items or they are not readily available because of time constraints in the pesticide review process. Further, the previous three areas concerned residue data collected and submitted by the pesticide registrant. For this area the registrant may or may not, review the literature; it is not a requirement, but it is recommended.

Summary

The ecological risk assessment usually utilizes the residue data generated with the chemical fate and transport data since these residues are the most readily available. Such data are used to determine the fate and transport of the pesticide. Unfortunately, said data

are of limited value relative to wildlife food items. To address residues on wildlife food, the Agency develops estimates of residue levels. A discussion of how this is done is presented below.

Estimated Residues

The estimated acute terrestrial residues utilized by the Agency are those shown in Table E-2 and they are based primarily upon the works of Hoerger and Kenaga (1972) and Kenaga (1973). This residue profile provides a worst case scenario, that is, the maximum expected residues likely to be found in or on vegetation and/or invertebrate (insect) surfaces immediately after application. This approach maximizes acute hazard determinations because day-zero (the day of application) residues are utilized.

Table E-2.

Maximum Expected Residues and Typical Residues of Pesticides on Differing Categories of Vegetation Types (from Hoerger and Kenaga, 1972).

	Residues (in ppm) for a Pesticide Dosage of 1 Lb/Acre			
	Immediately After Application		Six Weeks After Application	
Plant Category	Upper Limit	Typical Limit	Upper Limit	Typical Limit
Range Grass	240	125	30	5
Grass	110	92	20	1-5
Leaves and Leafy Crops	125	35	20	<1
Forage Crops (Small Insects)	58	33	1	<1
Pods Containing Seeds (Large Insects)	12	3	1.5	<1
Grain (Large Insects)	10	3	1.5	<1
Fruit (Large Insects)	7	1.5	1.5	<0.2

The Agency considers this approach reasonable because; 1) in most instances actual residue data are lacking, 2) the data presented by Hoerger and Kenaga (1972) appear to correlate fairly well with those of other researchers, 3) the pesticides and crops considered by the authors cover those reviewed routinely by the Agency and 4) as mentioned earlier, the Agency makes every attempt to correlate these estimates with actual residue data on pertinent wildlife food items (Urban and Cook, 1986).

In using this approach, the EPA:

- Uses the extrapolation proposed by Kenaga (1973) which is that residues on insects can be estimated from residue data for plants, or plant parts, with the same surface area to mass ratio as the insects in question. For small insects the values for dense foliage (forage crops) are used; for large insects the values for pods, grain and even fruit can be utilized;

- Utilizes a quantity of pesticide per square foot approach for assessing the risks of baits, seeds or granules to nontarget organisms (e.g., mg product/sq ft or number of seeds, baits or granules/sq ft); and

- Recognizes that the upper limit residue values presented by Hoerger and Kenaga (1972) are in terms of wet weight whereas most bird consumption values for the avian dietary LC_{50} studies are presented in terms of dry weight. (When appropriate and logical, the EPA uses a factor of three for adjusting a dry-weight diet to an estimated wet-weight diet (e.g., LC_{50} + 3) as suggested by various authors.)

For chronic residues, the EPA would correlate the available chemical fate and transport data with the acute terrestrial EECs in an attempt to obtain decline curves for appropriate wildlife food items. Whenever possible, however, actual pesticide residue data would be utilized.

TOXICOLOGICAL HAZARD DATA

General

The toxicological or effects data utilized in the hazard portion of the terrestrial risk equation consist not only of the terrestrial toxicity data outlined previously, but also other data that prove to be pertinent to the Agency's terrestrial risk assessment. Such other data include freshwater (and, depending on the pesticide use pattern, marine/estuarine) vertebrate and invertebrate toxicity data, nontarget beneficial insect effects data and nontarget plant effects data. These data can be acute, subacute and/or long term or chronic in nature; in most cases the data are developed in laboratory studies. For a full listing of the types of toxicity or effects data that can be required by the Agency the following should be consulted:

- The various subdivisions of the Pesticide Assessment Guidelines; and

- 40 CFR (158), Data Requirements for Pesticide Registration; Final Rule; Wednesday, October 24, 1984.

It should be noted that the majority, if not almost all, of the toxicity/effects data utilized by the EPA are developed and submitted by the pesticide registrant. Consequently, the registrant has certain statutory rights concerning citation of these studies and they cannot be used by other applicants without permission from or offering compensation to the owner.

Study Reliability/Statistics

Without going into the actual study design of the different toxicity/effects studies, it can be said that each study, whether short term (some are performed in 48 hours) or long term and highly complex (some are performed over weeks, months or even years), is critically reviewed by the Agency's scientific staff. The study's acceptability, relative to good scientific practice and its ability to support the pesticide submission under Agency consideration, are determined. Further, each study receives a statistical evaluation and, typically, only those data with the best statistical reliability are used in the ecological risk assessment process. (Note that, if studies are determined to be totally unacceptable, they are not used. Marginal studies may be used, but they are identified as such.) Also, the Agency has developed and continues to develop, Standard Evaluation Procedures (SEPs) for each kind of data that is required to support a pesticide submission. These SEPs

present the procedures used to evaluate the toxicity/effects data submitted and ensure that comprehensive and consistent treatment of the science in the data reviews is maintained by Agency staff (Urban and Cook, 1986).

INTEGRATION OF EXPOSURE AND TOXICOLOGICAL HAZARD DATA

General

Obviously, the critical step in any risk assessment is the integration of exposure and toxicological hazard data into a statement or conclusion concerning the risks to those organisms of concern when exposed to the items under study (in this case, pesticides). Generally, the reliability of the risk assessment is greater when the statistical reliability and scientific accuracy of the available data is high. For non-human, nontarget organisms the Agency makes every attempt to achieve such a desirable scenario. Unfortunately this situation is often not obtained because, typically, the Agency only has available:

- A toxicity/effects data set that does not contain all of the required terrestrial studies. These studies might include: one avian acute oral LD_{50} study, two avian dietary LC_{50} studies and acute, subacute and/or chronic studies with domestic mammals. Two avian reproduction studies may be required in some use situations.

- A limited number of test species used in the laboratory studies: e.g., mallard duck, bobwhite quail, ring-necked pheasant, rat, mouse, dog, guinea pig and rabbit.

- A limited number of data points. Generally, only five or six dose/concentration levels are used in the acute studies to develop the LD_{50} or LC_{50}.

- Laboratory results from lower-tiered terrestrial effects studies that are difficult to extrapolate to many field situations.

- Estimated Environmental Concentrations (EECs) rather than actual field residue data for pesticides.

From this it can be seen that the Agency is often extrapolating from a situation of limited information to a "real world" that has multiple species, animal populations and endangered species that are sensitive to ecological perturbations. To perform such extrapolations, a link is needed between the observed laboratory effects (or pharmacological vulnerability) and the estimated field effects (or ecological vulnerability) (Hudson *et al.*, 1984). Terrestrial field studies serve as the link. They are studies designed to determine what effects, if any, occur under actual pesticide use conditions. In essence, the results of such studies either support or refute the Agency's estimates of field effects.

Specific Extrapolative Techniques

The actual integration process requires the Agency to carefully correlate the exposure and toxicological hazard data discussed above. Terrestrial EECs, actual pesticide residue data, the pesticide's physical and chemical properties, pesticide use information, chemical fate and transport data, and nontarget organism/habitat information are integrated with the available laboratory (and, possibly, field or pesticide incidents data) mammalian, avian

and aquatic effects/ toxicity data. Determinations on potential acute, subacute, secondary and/or chronic risks are developed for both nonendangered and endangered nontarget fish and wildlife. Further, conclusions concerning: labeling, use restrictions, classification of uses (e.g., Restricted Use), the need for a Special Review, whether the product should be registered or not and the need for further data (e.g., terrestrial field study) are made. A complete discussion of the Agency's extrapolative techniques for determining what field effects are likely based on effects observed in laboratory studies and using aquatic and terrestrial EECs is presented in the EPA's SEP, Ecological Risk Assessment, EPA540/ 9-85-001, June, 1986. Specifically, the terrestrial risk assessment procedures are presented on pages 29 through 52. For convenience and brevity those techniques will not be repeated here, but interested parties should consult that document.

Dose-Response Curves

An especially critical part of the toxicological hazard data set the Agency uses in its ecological risk assessment and for determining when terrestrial field studies are required is the dose-response curves developed for LD_{50}/LC_{50} studies. In utilizing these curves, the EPA critically examines the study design of each LD_{50}/LC_{50} study and performs a variety of functions including: 1) recalculating and verifying the statistical results (using, for example, Finney Probit), 2) examining closely the variability of the test results, particularly the 95 percent confidence limits for the LD_{50}/LC_{50} values, 3) examining the observed and expected results closely at the 100 percent, 50 percent and 0 percent response levels and at the lowest effect level (LEL), 4) checking the slope of the dose-response curve and 5) noting the toxic symptoms and any sublethal responses that occur during the study (AIBS, 1978). LD/LC values other than the LD_{50}/LC_{50} may be developed, but with the knowledge that: the most statistically precise value is the LD_{50}/LC_{50} value; such extreme values as LD_{10}/ LC_{10} or LD_{99}/LC_{99} may not be accurate due to curvature of the dose-response line; and specially designed studies are actually needed to determine accurately such extreme values (Heath *et al.*, 1972; Hill *et al.*, 1975; Hudson *et al.*, 1984).

The Agency also critically examines longer term dose-response curves in a similar manner. At most, however, only three data points are available: a no effect level (NEL), a low effect level (LEL) and a high effect level (HEL).

"TRIGGERS" THAT REQUIRE FIELD STUDIES

There are several specific conditions or criteria that "trigger" terrestrial field studies requirement. Considering the above discussion, it can be seen that a flexible, weight-of-evidence approach is used by the Agency to perform an ecological risk assessment and to determine when terrestrial field studies are required. Many factors must be considered and integrated in the process. However, it is still possible to identify those conditions or criteria that must be met in order for a terrestrial field study to be required for a pesticide proposed for a particular use pattern.

Toxicity

The pesticide is acutely or chronically toxic to birds or mammals as shown by laboratory effects/toxicity studies. "Acutely" and "chronically" toxic are obviously relative terms to be used with discretion. Generally speaking, certain classes of compounds are considered to be highly toxic to groups of organisms whereas others are not. For example, many organophosphates and carbamates are considered to be acutely toxic to avian organisms. The Agency recognizes these chemical characteristics and, therefore,

uses the terms "acutely" and "chronically" in a flexible manner with the belief that specific criteria to distinguish between "acute" and "chronic" are inappropriate.

Exposure

Nontarget terrestrial organisms are likely to be exposed acutely and/or chronically to pesticide residues under actual field conditions. Specifically, nontarget organism, primarily avian and mammalian species, must be present in or adjacent to the treated areas. In essence, the likelihood of exposure for these organisms must be high.

EECs

Actual or estimated (terrestrial EECs) pesticide residues are present in the nontarget terrestrial organism's environment and are available to terrestrial organisms at levels equal to or greater than the acute and/or chronic lowest effect levels (LELs) observed in the laboratory effects/toxicity studies for birds and mammals.

Again, the Agency recognizes the limitations of using estimated residues, but does so only when pertinent actual residue data are lacking. Relative to use of the LEL, the EPA notes that other criteria (e.g., 1/5th or 1/10th of the LD_{50} or LC_{50}) have been used in the past in ecological risk assessments. However, in an attempt to be flexible and hopefully to include chemicals that might be of potential concern, the Agency has chosen to use the LEL. An extreme example, but one that supports this approach, would be a pesticide that caused blindness in test birds or mammals at the LEL. Obviously, the Agency would be seriously concerned with the potential risks to birds and mammals in the wild.

Residues

When the amount or duration of pesticide residues (as described above) increases relative to the acute and/or chronic effect levels observed in the laboratory effects/toxicity studies for mammals and birds, the Agency's ecological concerns increase and the likelihood that a terrestrial field study is required increases also.

This criterion is more open-ended, but it correlates exposure data on actual or estimated residues with toxicological hazard data.

Acute Risks or Concerns- Although the EPA has no specific "cut-off" point (for an EEC, actual residue value or toxicological effect) that can be presented here, the following can be stated:

Residue/Effect Level	Is Terrestrial Field Study Required?
Residue < NEL	No
NEL < Residue < LEL	No
LEL < Residue < LD_{50}/LC_{50}	Possibly
LD_{50}/LC_{50} < Residue	Yes

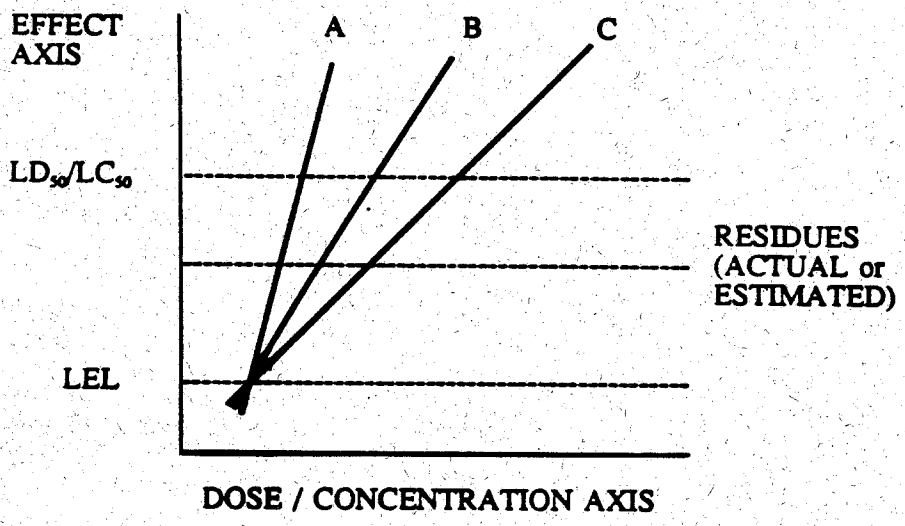
Generally speaking, the Agency has minimal concerns when actual or estimated pesticide residues (that are acute or of short duration) are below the LEL (as determined in laboratory studies). As these residues increase relative to the LD_{50} or LC_{50} values determined in laboratory studies, the Agency's ecological concerns increase and the likelihood of requiring terrestrial field studies increases.

In utilizing these ratios of residues to effect levels, the Agency must closely examine the acute dose-response curves developed in the laboratory acute effects/toxicity studies for mammals and birds. As an example, the following three hypothetical acute dose-response curves for A, B and C are presented to clarify the above and, particularly, the situation when: LEL < Residue < LD_{50}/LC_{50}: In Figure E-2:

- Each dose-response line has the same LEL; but
- Each dose-response curve (adjusted to a straightline via logs and probits) has a different LD_{50} or LC_{50}; and
- Each line has a different probit/log cycle or slope.

For the sake of discussion, each line crosses the LEL at the same point (obviously, a highly unlikely situation), but the increases in numbers affected (in this case by mortality) relative to dose/concentration are markedly different.

Figure E-2.



For the dose-response line (A) even small increases in dose/concentration result in significant increases in effect such that the LD_{50}/LC_{50} is quickly reached. For line (B) the increases are more gradual and for line (C) they are even more gradual. Essentially then, greater increases in dose or concentration are required to produce increased mortality and to reach the LD_{50} or LC_{50} for lines (B) and (C). The point of this discussion is to show that, if all other conditions are equal for (A), (B) and (C) (e.g., application rates, use sites, etc.), then, for acute risks, the dose-response becomes critical in determining when terrestrial field studies are required. Actual or estimated pesticide residues lying somewhere between the LEL and LD_{50} or LC_{50} (see example, as shown) would increase the Agency's ecological concerns and the likelihood that the terrestrial field study is required, more for line (B) than (C) and even more for (A) than either (B) or (C).

Chronic Risks or Concerns

For chronic risks or concerns a terrestrial field study is generally required when the pesticide residues (actual or estimated) equal or exceed the LEL. Because the LEL is usually an effect on a reproductive parameter and the potential for adverse population effects can be greater for chronic risks than for acute risks, the Agency is more conservative and requires the terrestrial field study in order to address the potential chronic or reproductive risks. Also, mitigation of chronic risks by label use restrictions and/or restricted use classification may not be as readily achieved as for acute risks.

Residue / Effect Level	Is Terrestrial Field Study Required?
Residue < NEL	No
NEL < Residue < LEL	No
LEL < Residue < HEL	Yes
HEL < Residue	Yes

SUMMARY

A variety of factors must be considered in performing an ecological risk assessment and in determining when a terrestrial field study is required. A flexible, weight-of-evidence approach is utilized in this process that can be represented by the relationship:

Toxicological Hazard Data	X	Exposure Data	=	Ecological Risks

However, four specific conditions or criteria (which, when met, "trigger" the requirement for a terrestrial field study), have been presented and discussed. These criteria are designed to be flexible and used with common sense and scientific integrity. Further, it is intended that said criteria will handle nearly all pesticide use situations, particularly those likely to impact on nontarget wildlife such as birds and mammals. It is also the Agency's intention that said criteria, the ecological risk assessment process and the terrestrial field study will address the potential risks, both acute and chronic, to terrestrial amphibians, reptiles and endangered organisms.

APPENDIX F

PAIRED PLOT DESIGN
BASIS FOR FORMULA FOR n PAIRS

This section presents, in detail, the basis for the formula to determine the number of blocks (pairs of plots) necessary to test the hypothesis that a pesticide has no short-term effect on wildlife mortality.

Suppose we have n blocks, each with a treated plot and an untreated (control) plot. Assume that the "true" number of "individuals" on the i^{th} control plot is C_i, and that the "true" number of "individuals" on the i^{th} treatment plot is T_i, for i = 1, 2, ... , n. Also, assume that $T_i = q\ C_i$ for each block, but that C_i need not be constant for all blocks. The parameter q is the survival ratio, with its complement, p = 1-q, being the short-term mortality ratio. For the totality of blocks, define parameters:

$$\overline{C} = (\Sigma\ C_i) \div n \quad \text{and} \quad \overline{T} = (\Sigma\ T_i) \div n\ ,$$

the means of the abundance (density) parameters C_i and T_i, respectively. Clearly,

$$\overline{T} = q\overline{C}.$$

Next, postulate that an observed count of the number of individuals on any plot is subject to measurement error. That is, if c_i and t_i are the observations for the i^{th} block, each is a single value from an infinite number of repeated, independent attempts to count the number of individuals. Assume that possible values of c_i and t_i follow independent Poisson distributions such that the probability of observing a specific value for c_i (or t_i) is a function of the abundance parameter C_i (or T_i) for the plot[1]. We may say that the variation of count for a plot is "locally Poisson." The mean and variance of a Poisson distribution are equal in value. For the totality of blocks, define the following statistics:

$$c = \Sigma\ c_i\ , \quad t = \Sigma\ t_i\ , \quad \overline{c} = c \div n, \quad \text{and} \quad \overline{t} = t \div n\ .$$

From distribution theory, it is known that the distribution of a sum of independent observations from different Poisson distributions is also a Poisson distribution with a parameter equal to the sum of the parameters of the distributions whose observations are summed. Proof of this assertion is available in Kendall and Stuart (1977, pp. 280-281), and need not be presented here. Hence, c has a Poisson distribution with parameter $n\overline{C}$, t has one with a parameter of $nq\overline{C}$, and w = c + t has one with a parameter $n\overline{C}(1+q)$.

At this point, we examine the conditional distribution of t, and also c, given a value for w = c + t. Note that c, t, and w each have discrete distributions. Denote the following probabilities:

P (c) - the probability of the event (value of) c,

P (t) - the probability of the event (value of) t,

---

[1] If y has a Poisson distribution with a parameter τ, then the probability that y is equal to a value r may be expressed as

$$P\ (y=r) = e^{-\tau}\ \ \tau^r \div r!$$

In this formula, e is the base of the natural logarithms, and r! is the factorial of r.

P (w) - the probability of the event (value of) w,

P (c,t) - the probability of the event *c and t*, and

P (t|w) - the probability of event t, given event w.

P (c,t) is known as a joint probability, while P (t|w) is a conditional probability.

Because of the assumption of independence of c_i and t_i, events c and t are also independent. Hence, P (c,t) = P (c) · P (t). By definition of a conditional probability,

$$P(t|w) = P(t,w) \div P(w).$$

Since $P(t,w) = P(c,t)$,

then $P(t|w) = P(c,t) \div P(w)$.

By substitution, we arrive at an intermediate result

$$P(t|w) = \frac{q^t}{(1+q)^w} \cdot \frac{w!}{c!\,t!}.$$

This may be rewritten as

$$P(t|w) = \frac{w!}{c!\,t!}\left(\frac{q}{1+q}\right)^t\left(\frac{1}{1+q}\right)^{w-t},$$

an expression that is readily recognized as the probability function for a binomial distribution with parameters w and P = q ÷ (1+q). We rewrite the equation as

$$P(t|w) = \binom{w}{t} P^t (1-P)^{w-t},$$

and note that w-t is c.

Therefore, the conditional distribution of t, given w is a binomial distribution. The limiting form of the binomial distribution with P near .5 and large w is a normal distribution with mean equal to wP, and variance equal to wP(1-P). This result suggests that, by defining $\hat{P} = t \div w$, we may obtain an estimator of P that is normally distributed with mean P and variance equal to P(1-P) ÷ w. This merely represents a simple linear transformation of the conditional binomial distribution of t.

Suppose we test the null hypothesis H_0: $q=q_0$ with a level of significance equal to α, against an alternative hypothesis H_1: $q=q_1$, with power equal to 1-ß, for $q_1 < q_0$. Equivalent hypotheses are H_0: $P=P_0$, and H_1: $P=P_1$, with $P_0 = q_0 \div (1+q_0)$ and $P_1 = q_1 \div (1+q_1)$. A critical value of $\hat{P}$, may be designated as $\hat{P}_d$ to represent the point on the scale of $\hat{P}$ that divides the scale into two decision regions; values of $\hat{P} < \hat{P}_d$ correspond to a decision to reject H_0, those where $\hat{P} > \hat{P}_d$ to nonrejection of H_0.

Now, $Z_\alpha = (\hat{P}_d - P_0) \div \sqrt{P_0 (1-P_0) \div w}$ and

$$Z_{1-\beta} = (\hat{P}_d - P_1) \div \sqrt{P_1 (1-P_1) \div w}$$

are values of $\hat{P}_d$ transformed to Z-scores with respect to the distribution of $\hat{P}$ under the null and alternative hypotheses. In order to simplify our algebra, we replace $P_0(1-P_0)$ by $P_1(1-P_1)$. This replacement makes little difference for null and alternative values of P relatively close to one another, and/or when w is relatively large with respect to $P_0(1-P_0)$.

Solving each of these equations for $\hat{P}_d$, and setting the results equal to each other, produces an equation for w, written as

$$w = (Z_{1-\alpha} + Z_{1-\beta})^2 \quad \frac{P_1 (1-P_1)}{(P_0-P_1)^2}$$

Since $w = c + t = n (\bar{c} + \bar{t})$, and on the average, $\bar{c} = \bar{C}$, and $\bar{t} = \bar{T} = q \bar{C}$, we replace w by $n \bar{c} (1+q_1)$. The choice of q_1 results in a slightly larger value of n, than if q_0 is selected. Also, we substitute the following quantities:

$$P_0 = 1/2 \text{ (when } q_0=1), \quad P_1 (1-P_1) = q_1 \div (1+q_1)^2, \text{ and}$$

$$(P_0-P_1)^2 = \frac{P_1^2}{4 (1+q_1)^2}$$

Finally,

$$n = (Z_{1-\alpha} + Z_{1-\beta})^2 \quad \frac{4q_1}{p_1^2 (1+q_1) \bar{c}}$$